

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:30:32 ; Search time 19 Seconds
(without alignments)
176.615 Million cell updates/sec

Title: US-09-955-502-1

Perfect score: 139

Sequence: 1 MRRXXCXKXXXXXXXXXXXX.....QTLXNEXLXXXXXXXXXX 65

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database :

Issued Patents AA.*

- 1: /cgm2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgm2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgm2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgm2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgm2_6/ptodata/2/iaa/PTCUS_COMB.pep.*
- 6: /cgm2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	34.5	93	4	US-09-543-681A-5443
2	48	34.5	107	4	US-09-489-039A-11962
3	47	33.8	110	4	US-09-540-236-2859
4	46	33.1	92	4	US-09-328-352-5456
5	44	31.7	122	4	US-09-252-991A-23355
6	34	24.5	180	4	US-09-134-000C-6419
7	34	24.5	356	4	US-08-178-257-5
8	34	24.5	589	2	US-08-756-317-5
9	33	23.7	571	4	US-09-489-039A-14334
10	31	22.3	214	4	US-09-543-681A-4972
11	30	21.6	285	4	US-09-252-991A-18318
12	29	20.9	131	4	US-09-213-293D-15
13	29	20.9	170	4	US-09-489-039A-7536
14	29	20.9	483	4	US-09-148-545-144
15	29	20.9	488	4	US-09-148-545-209
16	29	20.9	488	4	US-09-339-159B-14
17	29	20.9	511	4	US-09-134-000C-5362
18	29	20.9	519	4	US-09-252-991A-19734
19	29	20.9	664	3	US-09-295-186-17
20	29	20.9	778	3	US-09-460-145-2
21	29	20.9	778	3	US-09-895-547-2
22	29	20.9	797	3	US-09-460-145-4
23	29	20.9	797	4	US-09-895-547-4
24	29	20.9	913	3	US-08-827-208-3
25	29	20.9	913	3	US-09-500-358-3
26	29	20.9	913	3	US-09-498-809-3
27	28	20.1	23	1	US-07-646-531D-5

ALIGNMENTS

RESULT 1

US-09-543-681A-5443
; Sequence 5443 Application US/09543681A
; Patent No. 6805709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709,1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5443
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5443

Query Match 34.5%; Score 48; DB 4; Length 93;
Best Local Similarity 23.6%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 MRRXXCXKXXXXXXXXXXXXXGXXXXXXXXXXXXXXXXXXXXXXXXXXXXXQTLXNEXL 55
DB 4 MSRTIFTLNKEADGLDFQLYFGLKGRIFNEISKEAWGQWMAQTWLINEKKL 58

RESULT 2

US-09-489-039A-11962
; Sequence 11962, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11962
; LENGTH: 107

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/178,257
APPLICATION NUMBER: US/08/178,257
FILING DATE: 11-JAN-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/01291
FILING DATE: 15-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9115245.4
FILING DATE: 16-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 9320/206058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/861-3000
TELEFAX: 202/822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Alcaligenes eutrophus
US-08-178-257-5

Query Match 24.5%; Score 34; DB 4; Length 356;
Best Local Similarity 42.9%; Pred. No. 5.4;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 264 WYCWYLRHTYLQNE 277

RESULT 8
US-08-756-317-5
Sequence 5, Application US/08756317
Patent No. 5849894
GENERAL INFORMATION:
APPLICANT: Clemente, Thomas B.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Mitsky, Timothy A.
APPLICANT: Stark, David M.
TITLE OF INVENTION: Improved Rhodospirillum Rubrum
TITLE OF INVENTION: Poly-B-Hydroxyalkonocate Synthase
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,317
FILING DATE: 25-NOV-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,693
FILING DATE: 29-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Patterson, Melinda L.

REGISTRATION NUMBER: 33,062
REFERENCE/DOCKET NUMBER: MOBT:008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 787-1400
TELEFAX: (713) 787-1440
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 589 amino acids
TYPE: amino acid
STRADEDNESS:
TOPOLOGY: linear
US-08-756-317-5

Query Match 24.5%; Score 34; DB 2; Length 589;
Best Local Similarity 42.9%; Pred. No. 8.7;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 436 WYCWYLRHTYLQNE 449

RESULT 9
US-09-489-039A-14334
Sequence 14334, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 14334
LENGTH: 571
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14334

Query Match 23.7%; Score 33; DB 4; Length 571;
Best Local Similarity 35.3%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
DB 403 WQAWHDALTELLNRGAL 419

RESULT 10
US-09-543-681A-4972
Sequence 4972, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 4972
LENGTH: 214
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-4972

Query Match 22.3%; Score 31; DB 4; Length 214;
Best Local Similarity 29.4%; Pred. No. 16;

Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 89 WYGNWAFKTAFTKEIEL 105

RESULT 11

US-09-252-991A-18318
; Sequence 18318, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18318
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18318

Query Match 21.6%; Score 30; DB 4; Length 285;
Best Local Similarity 28.6%; Pred. No. 36;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNE 52
| | | | |
Db 56 WRSWVTTRNEVCNE 69

RESULT 12

US-09-213-293D-15
; Sequence 15, Application US/09213293D
; Patent No. 6384299
; GENERAL INFORMATION:
; APPLICANT: GUTIERREZ-ARMENTA, CRISANTO
; APPLICANT: SANZ-BURGOS, ANDRES P.
; APPLICANT: XIE, QI
; APPLICANT: LOPEZ, PAULA S.
; TITLE OF INVENTION: PLANT RETINOBLASTOMA-ASSOCIATED PROTEINS
; FILE REFERENCE: 604-469
; CURRENT APPLICATION NUMBER: US/09/213,293D
; CURRENT FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: PCT/ES96/00130
; PRIOR FILING DATE: 1996-06-13
; PRIOR APPLICATION NUMBER: PCT/EP97/03070
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Gallus sp.
US-09-213-293D-15

Query Match 20.9%; Score 29; DB 4; Length 131;
Best Local Similarity 42.9%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 42 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 43 WTLFQHTLQNESEL 56

RESULT 13

US-09-489-039A-7536
; Sequence 7536, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLBBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7536
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7536

Query Match 20.9%; Score 29; DB 4; Length 170;
Best Local Similarity 35.7%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNE 52
| | | | |
Db 64 WQLWLTPOQKLSRE 77

RESULT 14

US-09-148-545-144
; Sequence 144, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503

[illegible]

```
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 144
; LENGTH: 483

Query Match      20.9%; Score 29; DB 4; Length 483;
Best Local Similarity 35.7%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXTXNXE 52
Db 224 WDRWRNQNLDKE 237

RESULT 15
US-09-148-545-209
; Sequence 209; Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
```

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 05-Sep-1997
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22

NUMBER OF SEQ ID NOS: 280
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 209
LENGTH: 483

Query Match 20.9%; Score 29; DB 4; Length 483;
Best Local Similarity 35.7%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTLXNE 52
DB 224 WDRWVQNANLDKE 237

RESULT 16

US-09-339-159B-14
Sequence 14, Application US/09339159B
Patent No. 6566114
GENERAL INFORMATION:
APPLICANT: Kauppinen, Markus
APPLICANT: Schulein, Martin
APPLICANT: Schmorri, Kirk
APPLICANT: Andersen, Lene
APPLICANT: Bjornvad, Mads
TITLE OF INVENTION: No. 6566114el Mannanases
FILE REFERENCE: 5440.204-US
CURRENT APPLICATION NUMBER: US/09/339,159B
CURRENT FILING DATE: 1999-06-24
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 488
TYPE: PRT
ORGANISM: Humicola insolens
US-09-339-159B-14

Query Match 20.9%; Score 29; DB 4; Length 488;
Best Local Similarity 44.4%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQQT 47
DB 212 WANWIAQQT 220

RESULT 17

US-09-134-000C-5362
Sequence 5362, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5362
LENGTH: 511
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-5362

Query Match 20.9%; Score 29; DB 4; Length 511;
Best Local Similarity 35.7%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTLXNE 52
DB 484 WICWKEYETKEKE 497

RESULT 18

US-09-252-991A-19734
; Sequence 19734, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107195.136
; CURRENT FILING DATE: 1998-02-18
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19734
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19734

Query Match 20.9%; Score 29; DB 4; Length 519;
Best Local Similarity 20.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 23 PXXGXXXXXXXGXXXXXXXGXXXXQT 47
DB 477 PSWAGASSNCSRHRSRWSWATAT 501

RESULT 19

US-09-295-186-17
; Sequence 17, Application US/09295186B
; Patent No. 6127137
; GENERAL INFORMATION:
; APPLICANT: Hasida, Miyoko
; APPLICANT: Tsutsumi, No. 6127137iko
; APPLICANT: Halkier, Torben
; APPLICANT: Stringer, Mary Ann
; TITLE OF INVENTION: An Acidic Phospholipase, Production, and
; FILE REFERENCE: 4953.204-US
; CURRENT APPLICATION NUMBER: US/09/295,186B
; CURRENT FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 1215/96
; PRIOR FILING DATE: 1996-10-31
; PRIOR APPLICATION NUMBER: PCT/DK97/00490
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 664
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-295-186-17

Query Match 20.9%; Score 29; DB 3; Length 664;
Best Local Similarity 30.8%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
DB 158 WNNWTSVQAIVDN 170

RESULT 20

US-09-460-145-2
; Sequence 2, Application US/09460145
; Patent No. 6287838
; GENERAL INFORMATION:

APPLICANT: Kriz, Ron
APPLICANT: Song, Chuanzheng
TITLE OF INVENTION: CYTOSOLIC PHOSPHOLIPASE A2-BETA ENZYMES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 97 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/460,145
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/788,975
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15289
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 778 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-460-145-2

Query Match 20.9%; Score 29; DB 3; Length 778;
Best Local Similarity 35.7%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 520 WDRWVRNQANLDKE 533

RESULT 21

US-09-895-547-2
; Sequence 2, Application US/09895547
; Patent No. 6482625
; GENERAL INFORMATION:
; APPLICANT: Kriz, Ron
; APPLICANT: Song, Chuanzheng
; TITLE OF INVENTION: CYTOSOLIC PHOSPHOLIPASE A2-BETA ENZYMES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 97 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/895,547
FILING DATE: 29-Jun-2001
CLASSIFICATION: <Unknown>

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/460,145
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brown, Scott A.
;; REGISTRATION NUMBER: 32,724
;; REFERENCE/DOCKET NUMBER: G15289
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 498-8224
;; TELEFAX: (617) 876-5851
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 778 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: <Unknown>
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-895-547-2

Query Match 20.9%; Score 29; DB 4; Length 778;
Best Local Similarity 35.7%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXTXLXNE 52
DB 520 WDRWVRNQANLDKE 533

RESULT 22
US-09-460-145-4
; Sequence 4, Application US/09460145
; Patent No. 6287838
; GENERAL INFORMATION:
; APPLICANT: Kriz, Ron
; TITLE OF INVENTION: CYTOSOLIC PHOSPHOLIPASE A2-BETA ENZYMES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/460,145
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/460,145
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: G15289
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 797 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-460-145-4

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/460,145
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brown, Scott A.
;; REGISTRATION NUMBER: 32,724
;; REFERENCE/DOCKET NUMBER: G15289
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 498-8224
;; TELEFAX: (617) 876-5851
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 778 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: <Unknown>
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-895-547-4

Query Match 20.9%; Score 29; DB 3; Length 797;
Best Local Similarity 35.7%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXTXLXNE 52
DB 539 WDRWVRNQANLDKE 552

RESULT 23
US-09-895-547-4
; Sequence 4, Application US/09895547
; Patent No. 6482625
; GENERAL INFORMATION:
; APPLICANT: Kriz, Ron
; TITLE OF INVENTION: CYTOSOLIC PHOSPHOLIPASE A2-BETA ENZYMES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/895,547
; FILING DATE: 23-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/460,145
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: G15289
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 797 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-895-547-4

Query Match 20.9%; Score 29; DB 4; Length 797;
Best Local Similarity 35.7%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXTXLXNE 52
DB 539 WDRWVRNQANLDKE 552

RESULT 24
US-08-827-208-3
; Sequence 3, Application US/08827208
; Patent No. 6025178
; GENERAL INFORMATION:
; APPLICANT: Chiou, Xue-Chiou C.
; APPLICANT: Kramer, Ruth M.
; APPLICANT: Pickard, Richard T.
; APPLICANT: Sharp, John D.
; APPLICANT: Striffler, Beth A.

;/ TITLE OF INVENTION: HUMAN PHOSPHOLIPASE A2 AND RELATED
;/ TITLE OF INVENTION: NUCLEIC ACID COMPOUNDS
;/ NUMBER OF SEQUENCES: 4
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Eli Lilly and Company
;/ STREET: Lilly Corporate Center
;/ CITY: Indianapolis
;/ STATE: Indiana
;/ COUNTRY: United States of America
;/ ZIP: 46285
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: PatentIn Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US 08/827,208
;/ FILING DATE: 28-MAR-1997
;/ CLASSIFICATION: 435
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 60/014,608
;/ FILING DATE: 29-MAR-1996
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 60/041,264
;/ FILING DATE: 19-MAR-1997
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Gaylo, Paul J.
;/ REGISTRATION NUMBER: 36,808
;/ REFERENCE/DOCKET NUMBER: X-10610
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (317) 276-0756
;/ TELEFAX: (317) 276-3861
;/ INFORMATION FOR SEQ ID NO: 3:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 913 amino acids
;/ TYPE: amino acid
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: protein
;/ US-08-827-208-3

Query Match 20.9%; Score 29; DB 3; Length 913;
Best Local Similarity 35.7%; Pred. No. 1.8e-02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 655 WDRWVNQANLDKE 668

RESULT 25
US-09-500-358-3
;/ Sequence 3, Application US/09500358
;/ Patent No. 6197569
;/ GENERAL INFORMATION:
;/ APPLICANT: Chiou, Xue-Chiou C.
;/ APPLICANT: Kramer, Ruth M.
;/ APPLICANT: Pickard, Richard T.
;/ APPLICANT: Sharp, John D.
;/ APPLICANT: Striffler, Beth A.
;/ TITLE OF INVENTION: HUMAN PHOSPHOLIPASE A2 AND RELATED
;/ NUMBER OF SEQUENCES: 4
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Eli Lilly and Company
;/ STREET: Lilly Corporate Center
;/ CITY: Indianapolis
;/ STATE: Indiana
;/ COUNTRY: United States of America
;/ ZIP: 46285
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS

;/ SOFTWARE: PatentIn Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/09/500,358
;/ FILING DATE:
;/ CLASSIFICATION:
;/ PRIOR APPLICATION DATA: US/08/827,208
;/ APPLICATION NUMBER:
;/ FILING DATE: 28-MAR-1997
;/ APPLICATION NUMBER: US 60/014,608
;/ FILING DATE: 29-MAR-1996
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 60/041,264
;/ FILING DATE: 19-MAR-1997
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Gaylo, Paul J.
;/ REGISTRATION NUMBER: 36,808
;/ REFERENCE/DOCKET NUMBER: X-10610
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (317) 276-0756
;/ TELEFAX: (317) 276-3861
;/ INFORMATION FOR SEQ ID NO: 3:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 913 amino acids
;/ TYPE: amino acid
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: protein
;/ US-09-500-358-3

Query Match 20.9%; Score 29; DB 3; Length 913;
Best Local Similarity 35.7%; Pred. No. 1.8e-02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 655 WDRWVNQANLDKE 668

RESULT 26
US-09-498-809-3
;/ Sequence 3, Application US/09498809
;/ Patent No. 6242206
;/ GENERAL INFORMATION:
;/ APPLICANT: Chiou, Xue-Chiou C.
;/ APPLICANT: Kramer, Ruth M.
;/ APPLICANT: Pickard, Richard T.
;/ APPLICANT: Sharp, John D.
;/ APPLICANT: Striffler, Beth A.
;/ TITLE OF INVENTION: HUMAN PHOSPHOLIPASE A2 AND RELATED
;/ NUMBER OF SEQUENCES: 4
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Eli Lilly and Company
;/ STREET: Lilly Corporate Center
;/ CITY: Indianapolis
;/ STATE: Indiana
;/ COUNTRY: United States of America
;/ ZIP: 46285
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: PatentIn Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/09/498,809
;/ FILING DATE:
;/ CLASSIFICATION:
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: 08/827,208
;/ FILING DATE:
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 60/041,264
;/ FILING DATE: 19-MAR-1997
;/ ATTORNEY/AGENT INFORMATION:

NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-10610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 276-0756
TELEFAX: (317) 276-3861
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 913 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-498-809-3

Query Match 20.1%; Score 29; DB 3; Length 913;
Best Local Similarity 35.7%; Pred. No. 1.Be+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 39 WXXWXXXTXKNE 52
Db 655 WDRVNRQANLDKE 668

RESULT 27

US-07-646-531D-5
Sequence 5, Application US/07646531D
Patent No. 5200397
GENERAL INFORMATION:
APPLICANT: Deutch, Alan Howard
TITLE OF INVENTION: Peptide Fragments and Analogs of
Thrombospondin
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: W. R. Grace & Co.-Conn.
STREET: 7379 Route 32
CITY: Columbia
STATE: Maryland
COUNTRY: USA
ZIP: 21044
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/646,531D
FILING DATE: 19910131
ATTORNEY/AGENT INFORMATION:
NAME: Appleby, Vanessa L.
REGISTRATION NUMBER: 33223
REFERENCE/DOCKET NUMBER: 01-7896
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 531-4515
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-646-531D-5

Query Match 20.1%; Score 28; DB 1; Length 23;
Best Local Similarity 30.8%; Pred. No. 9.5;
Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 39 WXXWXXXTXKNE 51
Db 4 WSEWTSCTSCGN 16

RESULT 28

US-08-488-273-5

Sequence 5, Application US/08488273
Patent No. 5840692
GENERAL INFORMATION:
APPLICANT: Deutch, Alan H.
APPLICANT: Tuszyński, George P.
TITLE OF INVENTION: PEPTIDE FRAGMENTS AND ANALOGS OF
THROMBOSPONDIN
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,273
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/359,263
FILING DATE: 19-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,146
FILING DATE: 20-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/896,527
FILING DATE: 09-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/483,527
FILING DATE: 22-FEB-1990
ATTORNEY/AGENT INFORMATION:
NAME: Leary Ph.D., Kathryn
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-2U4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2020
TELEFAX: (215) 567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-273-5

Query Match 20.1%; Score 28; DB 2; Length 23;
Best Local Similarity 30.8%; Pred. No. 9.5;
Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 39 WXXWXXXTXKNE 51
Db 4 WSEWTSCTSCGN 16

RESULT 29

US-09-197-770B-11
Sequence 11, Application US/09197770B
Patent No. 6339062
GENERAL INFORMATION:
APPLICANT: Tuszyński, George
APPLICANT: Williams, Taffy
APPLICANT: Actor, Paul
TITLE OF INVENTION: RETROINVERSO POLYPEPTIDES THAT MIMIC OR INHIBIT
THROMBOSPONDIN ACTIVITY
FILE REFERENCE: 07206-0021

; CURRENT APPLICATION NUMBER: US/09/197,770B
; CURRENT FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: fragment/ analog of thrombospondin
US-09-197-770B-11

Query Match 20.1%; Score 28; DB 4; Length 23;
Best Local Similarity 30.8%; Pred. No. 9.5;
Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 39 WXXWXXQTXLXN 51
| | | | |
Db 4 WSEWTSCTSCGN 16

RESULT 30
5426100-5
; Patent No. 5426100
; APPLICANT: DEUTCH, ALAN H.; TUSZYNSKI, GEORGE
; TITLE OF INVENTION: PEPTIDE FRAGMENTS AND ANALOGS OF
; THROMBOSPONDIN
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,146
; FILING DATE: 20-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 896,527
; FILING DATE: 09-JUN-1992
; APPLICATION NUMBER: 483,527
; FILING DATE: 22-FEB-1990
; SEQ ID NO: 5
; LENGTH: 23

Query Match 20.1%; Score 28; DB 5; Length 23;
Best Local Similarity 30.8%; Pred. No. 9.5;
Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 39 WXXWXXQTXLXN 51
| | | | |
Db 4 WSEWTSCTSCGN 16

RESULT 31
US-08-152-721B-28
; Sequence 28, Application US/08152721B
; Patent No. 5962315
; GENERAL INFORMATION:
; APPLICANT: Livingston, David M.
; APPLICANT: Ewen, Mark E.
; TITLE OF INVENTION: DNA Encoding p107 Tumor Suppressor and
; TITLE OF INVENTION: Related Polypeptides
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHOATE, HALL & STEWART
; STREET: 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2891
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/152,721B
; FILING DATE: 15-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasternack Esq., Sam
; REGISTRATION NUMBER: 29,576
; REFERENCE/DOCKET NUMBER: 181411-011DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-5000
; TELEFAX: (617) 248-4000
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; IMMEDIATE SOURCE:
; CLONE: RB region 6
US-08-152-721B-28

Query Match 20.1%; Score 28; DB 2; Length 71;
Best Local Similarity 42.9%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 42 WXXXQTXLXNEXXL 55
| | | | |
Db 7 WTLFQHTLQNEYEL 20

RESULT 32
US-09-213-293D-17
; Sequence 17, Application US/09213293D
; Patent No. 6384299
; GENERAL INFORMATION:
; APPLICANT: GUTIERREZ-ARMENTA, CRISANTO
; APPLICANT: SANZ-BURGOS, ANDRES P.
; APPLICANT: XIE, QI
; APPLICANT: LOPEZ, PAULA S.
; TITLE OF INVENTION: PLANT RETINOBLASTOMA-ASSOCIATED PROTEINS
; FILE REFERENCE: 604-469
; CURRENT APPLICATION NUMBER: US/09/213,293D
; CURRENT FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: PCT/ES96/00130
; PRIOR FILING DATE: 1996-06-13
; PRIOR APPLICATION NUMBER: PCT/EP97/03070
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 17
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-293D-17

Query Match 20.1%; Score 28; DB 4; Length 130;
Best Local Similarity 42.9%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 42 WXXXQTXLXNEXXL 55
| | | | |
Db 42 WTLFQHTLQNEYEL 55

RESULT 33
US-09-213-293D-16
; Sequence 16, Application US/09213293D
; Patent No. 6384299
; GENERAL INFORMATION:
; APPLICANT: GUTIERREZ-ARMENTA, CRISANTO

```
; APPLICANT: SANZ-BURGOS, ANDRES P.
; APPLICANT: XIE, QI
; APPLICANT: LOPEZ, PAULA S.
; TITLE OF INVENTION: PLANT RETINOBLASTOMA-ASSOCIATED PROTEINS
; FILE REFERENCE: 604-469
; CURRENT APPLICATION NUMBER: US/09/213,293D
; CURRENT FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: PCT/ES96/00130
; PRIOR FILING DATE: 1996-06-13
; PRIOR APPLICATION NUMBER: PCT/EP97/03070
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Mus sp.
; US-09-213-293D-16

Query Match      20.1%; Score 28; DB 4; Length 131;
Best Local Similarity 42.9%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 42 WXXXQTXLXNEXXL 55
Db 43 WTLFQHTLQNEYEL 56

RESULT 34
US-08-874-832-18
; Sequence 18, Application US/08874832
; Patent No. 5872011
; GENERAL INFORMATION:
; APPLICANT: Burly, Stephen K.
; APPLICANT: Sonenberg, Nahum
; APPLICANT: Marcotrigiano, Joseph
; APPLICANT: Gingras, Anne-Claude
; TITLE OF INVENTION: A CRYSTAL OF A PROTEIN-LIGAND COMPLEX
; TITLE OF INVENTION: CONTAINING AN N-TERMINAL TRUNCATED eIF4E, AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,832
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 2261-1-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-955-502-1.1.top50.ra1

Query Match      20.1%; Score 28; DB 4; Length 178;
Best Local Similarity 44.4%; Pred. No. 65;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 39 WXXXWXXXT 47
Db 8 WTFWYDIQT 16

RESULT 35
US-09-097-233-18
; Sequence 18, Application US/09097233
; Patent No. 6020162
; GENERAL INFORMATION:
; APPLICANT: Burley, Stephen K.
; APPLICANT: Sonenberg, Nahum
; APPLICANT: Marcotrigiano, Joseph
; APPLICANT: Gingras, Anne-Claude
; TITLE OF INVENTION: A CRYSTAL OF A PROTEIN-LIGAND COMPLEX
; TITLE OF INVENTION: CONTAINING AN N-TERMINAL TRUNCATED eIF4E, AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,233
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 2261-1-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-097-233-18

Query Match      20.1%; Score 28; DB 3; Length 178;
Best Local Similarity 44.4%; Pred. No. 65;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 39 WXXXWXXXT 47
Db 8 WTFWYDIQT 16

RESULT 36
US-08-874-832-9
; Sequence 9, Application US/08874832
; Patent No. 5872011
; GENERAL INFORMATION:
; APPLICANT: Burly, Stephen K.
```

```

; APPLICANT: Sonenberg, Nahum
; APPLICANT: Marcotrigiano, Joseph
; APPLICANT: Gingras, Anne-Claude
; TITLE OF INVENTION: A CRYSTAL OF A PROTEIN-LIGAND COMPLEX
; TITLE OF INVENTION: CONTAINING AN N-TERMINAL TRUNCATED EIF4E, AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hacksack Ave, Continental Plaza, 4th
; CITY: Hacksack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,832
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 2261-1-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-874-832-9

Query Match 20.1%; Score 28; DB 2; Length 209;
Best Local Similarity 44.4%; Pred. No. 76;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
DB 39 WTFWYDIQT 47

RESULT 37
US-09-097-233-9
; Sequence 9, Application US/09097233
; Patent No. 6020182
; GENERAL INFORMATION:
; APPLICANT: Burley, Stephen K.
; APPLICANT: Sonenberg, Nahum
; APPLICANT: Marcotrigiano, Joseph
; APPLICANT: Gingras, Anne-Claude
; TITLE OF INVENTION: A CRYSTAL OF A PROTEIN-LIGAND COMPLEX
; TITLE OF INVENTION: CONTAINING AN N-TERMINAL TRUNCATED EIF4E, AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hacksack Ave, Continental Plaza, 4th
; CITY: Hacksack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-985-526-1

; APPLICANT: Sonenberg, Nahum
; APPLICANT: Marcotrigiano, Joseph
; APPLICANT: Gingras, Anne-Claude
; TITLE OF INVENTION: A CRYSTAL OF A PROTEIN-LIGAND COMPLEX
; TITLE OF INVENTION: CONTAINING AN N-TERMINAL TRUNCATED EIF4E, AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hacksack Ave, Continental Plaza, 4th
; CITY: Hacksack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,233
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 2261-1-001 N
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-097-233-9

Query Match 20.1%; Score 28; DB 3; Length 209;
Best Local Similarity 44.4%; Pred. No. 76;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
DB 39 WTFWYDIQT 47

RESULT 38
US-08-985-526-1
; Sequence 1, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutcz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-985-526-1
```

```
Query Match      20.1%; Score 28; DB 3; Length 218;
Best Local Similarity 30.8%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXN 51
Db 85 WSEWTSCTSCGN 97

RESULT 39
US-09-252-991A-22823
; Sequence 22823, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22823
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22823

Query Match      20.1%; Score 28; DB 4; Length 220;
Best Local Similarity 44.4%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTX 47
Db 44 WQWQKQYQTX 52

RESULT 40
US-09-134-000C-4928
; Sequence 4928, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4928
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4928

Query Match      20.1%; Score 28; DB 4; Length 234;
Best Local Similarity 29.4%; Pred. No. 84;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNXXL 55
Db 126 WAEWWEEMPRTEGGL 142

RESULT 41
PCT-US93-01652-1
```

```
; Sequence 1, Application PC/TUS9301652
; GENERAL INFORMATION:
; APPLICANT: Bouck, Noel P.
; APPLICANT: Polverini, Peter J.
; APPLICANT: Good, Deborah J.
; APPLICANT: Frazier, William A.
; TITLE OF INVENTION: Method and Composition for
; TITLE OF INVENTION: Inhibiting Angiogenesis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01652
; FILING DATE: 19930222
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/841,656
; FILING DATE: 24-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/464,369
; FILING DATE: 12-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fentress, Susan B.
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92005-PCT
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US93-01652-1

Query Match      20.1%; Score 28; DB 5; Length 239;
Best Local Similarity 30.8%; Pred. No. 86;
Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXN 51
Db 74 WSEWTSCTSCGN 86

RESULT 42
US-09-252-991A-32389
; Sequence 32389, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32389
```

```
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32389

Query Match      20.1%; Score 28; DB 4; Length 389;
Best Local Similarity 28.8%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
Db 345 WSYWDTDTTHLFS 358

RESULT 43
US-08-985-526-3
; Sequence 3, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-985-526-3

Query Match      20.1%; Score 28; DB 3; Length 441;
Best Local Similarity 30.8%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 308 WSEWTSCSTSCGN 320

RESULT 44
US-09-252-991A-26556
; Sequence 26556, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
```

```
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26556
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26556

Query Match      20.1%; Score 28; DB 4; Length 450;
Best Local Similarity 23.5%; Pred. No. 1.6e+02;
Matches 4; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXL 55
Db 219 WSRWRTSTSMESMTAL 235

RESULT 45
US-09-092-636-9
; Sequence 9, Application US/09092636A
; Patent No. 6162641
; GENERAL INFORMATION:
; APPLICANT: Goldman, Daniel
; APPLICANT: Sapru, Mohan K.
; TITLE OF INVENTION: Neuregulin Response Element and Uses Therefor
; FILE REFERENCE: UMI-003
; CURRENT APPLICATION NUMBER: US/09/092,636A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-092-636-9

Query Match      20.1%; Score 28; DB 3; Length 468;
Best Local Similarity 30.8%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 114 WLLWATNEFSLVN 126

RESULT 46
US-09-092-636-4
; Sequence 4, Application US/09092636A
; Patent No. 6162641
; GENERAL INFORMATION:
; APPLICANT: Goldman, Daniel
; APPLICANT: Sapru, Mohan K.
; TITLE OF INVENTION: Neuregulin Response Element and Uses Therefor
; FILE REFERENCE: UMI-003
; CURRENT APPLICATION NUMBER: US/09/092,636A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-092-636-4

Query Match      20.1%; Score 28; DB 3; Length 470;
Best Local Similarity 30.8%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
```


Qy 39 WXXWXXXQTXLXN 51
| | | | |
Db 114 WLHWATNEFSLVN 126

RESULT 47

US-09-489-039A-14269
; Sequence 14269, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 14269
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14269

Query Match 20.1%; Score 28; DB 4; Length 523;
Best Local Similarity 44.4%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTX 47
| | | | |
Db 247 WARWMEQT 255

RESULT 48

US-09-075-505-7
; Sequence 7, Application US/09075505
; Patent No. 6489305
; GENERAL INFORMATION:

; APPLICANT: Demers, G. William
; TITLE OF INVENTION: Methods and Compositions for the
; TITLE OF INVENTION: Treatment of Ocular Diseases
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,505
; FILING DATE: 08-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 016930-003600US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-075-505-7

Query Match 20.1%; Score 28; DB 4; Length 550;
Best Local Similarity 42.9%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 42 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 303 WTLFQHTLQNEVEL 316

RESULT 49

US-09-543-681A-6987
; Sequence 6987, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6987
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6987

Query Match 20.1%; Score 28; DB 4; Length 648;
Best Local Similarity 44.4%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTX 47
| | | | |
Db 391 WETWEEMQT 399

RESULT 50

US-09-295-186-18
; Sequence 18, Application US/09295186B
; Patent No. 6127137
; GENERAL INFORMATION:
; APPLICANT: Hasida, Miyoko
; APPLICANT: Tsutsumi, No. 6127137iko
; APPLICANT: Halkier, Torben
; APPLICANT: Stringer, Mary Ann
; TITLE OF INVENTION: An Acidic Phospholipase, Production, and
; TITLE OF INVENTION: Methods of Using Thereof (As Amended)
; FILE REFERENCE: 4953.204-US
; CURRENT APPLICATION NUMBER: US/09/295,186B
; CURRENT FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 1215/96
; PRIOR FILING DATE: 1996-10-31
; PRIOR APPLICATION NUMBER: PCT/DK97/00490
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Torulaspora delbrueckii
US-09-295-186-18

Query Match 20.1%; Score 28; DB 3; Length 649;
Best Local Similarity 30.8%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXN 51
| | | | |
Db 156 WNNWTSVQDIVNN 168

Thu Sep 30 14:24:05 2004

us-09-955-502-1.top50.ra

Page 18

Search completed: September 30, 2004, 10:33:19
Job time : 21 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:28:06 ; Search time 57 Seconds
(without alignments)
322.203 Million cell updates/sec

Title: US-09-955-502-1
Perfect score: 139
Sequence: 1 MRRXXXCXXXXXXX.....QTLXNEXLXXXXXXX 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : A_Geneseq_29Jan04.*

1: geneseqp1988s.*
2: geneseqp1998s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	35.3	90	ABB78165	Abb78165 Amino aci
2	48	34.5	78	ABB78164	Abb78164 Amino aci
3	48	34.5	88	ABB78178	Abb78178 Amino aci
4	48	34.5	88	ABB78160	Abb78160 Amino aci
5	48	34.5	91	ABB78161	Abb78161 Amino aci
6	48	34.5	91	ABB78158	Abb78158 Amino aci
7	48	34.5	91	ABB78159	Abb78159 Amino aci
8	48	34.5	91	ABB78162	Abb78162 Amino aci
9	48	34.5	91	ABB78163	Abb78163 Amino aci
10	48	34.5	91	ABB78157	Abb78157 Amino aci
11	48	34.5	91	ABB78156	Abb78156 Amino aci
12	47	33.8	87	ABB78151	Abb78151 Amino aci
13	47	33.8	87	ABB78152	Abb78152 Amino aci
14	47	33.8	87	ABB78153	Abb78153 Amino aci
15	47	33.8	90	ABB78155	Abb78155 Amino aci
16	46	33.1	76	ABB78166	Abb78166 Amino aci
17	46	33.1	88	ABB78171	Abb78171 Amino aci
18	46	33.1	88	ABB78172	Abb78172 Amino aci
19	46	33.1	88	ABB78154	Abb78154 Amino aci
20	46	33.1	88	ABB78173	Abb78173 Amino aci
21	46	33.1	88	ABP77219	ABP77219 N. Gonorr
22	46	33.1	90	ABB78168	Abb78168 Amino aci
23	46	33.1	91	ABB78150	Abb78150 Amino aci
24	46	33.1	92	ADA34169	ADA34169 Acinetoba
25	45	32.4	87	ABB78175	Abb78175 Amino aci

26	45	32.4	87	5	ABB78169	Abb78169 Amino aci
27	45	32.4	87	5	ABB78177	Abb78177 Amino aci
28	45	32.4	87	5	ABB78176	Abb78176 Amino aci
29	44	31.7	86	5	ABB78149	Abb78149 Amino aci
30	44	31.7	87	5	ABB78148	Abb78148 Amino aci
31	44	31.7	87	5	ABB78170	Abb78170 Amino aci
32	44	31.7	87	5	ABB78147	Abb78147 Amino aci
33	44	31.7	90	5	ABB78167	Abb78167 Amino aci
34	42	30.2	62	5	ABB78179	Abb78179 Consensus
35	39	28.1	87	5	ABB78174	Abb78174 Amino aci
36	35	25.2	1767	4	ABB70457	Abb70457 Drosophil
37	34	24.5	589	2	AAR10681	Aar10681 Polyhydro
38	34	24.5	589	2	AAR32190	Aar32190 Sequence
39	34	24.5	589	4	AAE10892	AAe10892 Alicaigen
40	34	24.5	589	5	ABU10991	Abu10991 Poly3-hyd
41	34	24.5	589	5	ABU10974	Abu10974 Poly3-hyd
42	34	24.5	589	5	ABU10975	Abu10975 Poly3-hyd
43	34	24.5	589	5	ABU10980	Abu10980 Poly3-hyd
44	34	24.5	589	5	ABU10978	Abu10978 Poly3-hyd
45	34	24.5	589	5	ABU10989	Abu10989 Poly3-hyd
46	34	24.5	589	5	ABU10962	Abu10962 Poly3-hyd
47	34	24.5	589	5	ABU10987	Abu10987 Poly3-hyd
48	34	24.5	589	5	ABU10981	Abu10981 Poly3-hyd
49	34	24.5	589	5	ABU10986	Abu10986 Poly3-hyd
50	34	24.5	589	5	ABU10988	Abu10988 Poly3-hyd

ALIGNMENTS

RESULT 1

ABB78165

ID ABB78165 standard; protein; 90 AA.

XX ABB78165;

DT 05-NOV-2002 (first entry)

DE Amino acid sequence of a YggX homologue.

XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
KW hydroxyl radical; DNA damage; YggX homologue.
XX Unidentified.

XX US2002072118-A1.

XX 13-JUN-2002.

XX 18-SEP-2001; 2001US-00955502.

XX 22-SEP-2000; 2000US-0234588P.

XX (DOWN/) DOWNS D.

XX (GRAL/) GRALNICK J A.

XX Downs D, Gralnick JA;

XX WPI; 2002-589476/63.

XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant cell, comprises engineering the cell to produce more YggX protein, a protein identified from Salmonella enterica Serovar Typhimurium.
XX Example; Fig 1A; 16pp; English.

XX The specification describes a method for reducing superoxide damage to a cell. The method comprises engineering the cell to produce more than the native amount of YggX protein (a protein identified from Salmonella enterica serovar typhimurium) or its homolog, where the cells are rendered more resistant to superoxide damage. YggX reduces the oxidation of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from clusters. The resulting decrease in free-iron levels generates fewer

CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues

XX
 SQ Sequence 90 AA;

Query Match 35.3%; Score 49; DB 5; Length 90;
 Best Local Similarity 23.6%; Pred. No. 0.0023;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXXCXNXXXXXXXXXXXXXGXXGXXXXXXXXXXXXXQTXLNEXXL 55
 Db 1 MSRTIFCTFLKDAERQDFOLYPGEIGKRIYNEISKEANSQWITKQTMLINEKKL 55

RESULT 2
 ABB78164
 ID ABB78164 standard; protein; 78 AA.
 XX
 AC ABB78164;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Amino acid sequence of a YggX homologue.
 XX
 KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 XX
 OS Klebsiella pneumoniae.
 XX
 PN US2002072118-A1.
 XX
 PD 13-JUN-2002.
 XX
 PF 18-SEP-2001; 2001US-00955502.
 XX
 PR 22-SEP-2000; 2000US-0234588P.
 XX
 PA (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 PI Downs D, Gralnick JA;
 XX
 DR WPI; 2002-589476/63.
 XX
 PT Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 XX
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues

XX
 SQ Sequence 78 AA;

Query Match 34.5%; Score 48; DB 5; Length 78;
 Best Local Similarity 23.6%; Pred. No. 0.0035;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXXCXNXXXXXXXXXXXXXGXXGXXXXXXXXXXXXXQTXLNEXXL 55
 Db 1 MSRTIFCTFLQREADGGDFOLYPGELGKRIYNEISKEANQWQHQTMLINEKKL 55

RESULT 3
 ABB78178
 ID ABB78178 standard; protein; 88 AA.
 XX
 AC ABB78178;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Amino acid sequence of a YggX homologue.
 XX
 KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 XX
 OS Coxiella burnetii.
 XX
 PN US2002072118-A1.
 XX
 PD 13-JUN-2002.
 XX
 PF 18-SEP-2001; 2001US-00955502.
 XX
 PR 22-SEP-2000; 2000US-0234588P.
 XX
 PA (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 PI Downs D, Gralnick JA;
 XX
 DR WPI; 2002-589476/63.
 XX
 PT Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 XX
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues

XX
 SQ Sequence 88 AA;

Query Match 34.5%; Score 48; DB 5; Length 88;
 Best Local Similarity 23.6%; Pred. No. 0.0039;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXXCXNXXXXXXXXXXXXXGXXGXXXXXXXXXXXXXQTXLNEXXL 55
 Db 1 MTRIIQKLGKADALNYSYPGELGERIYNHISEQAWQAWLSHOTMLINEYRL 55

RESULT 4
 ABB78160
 ID ABB78160 standard; protein; 88 AA.
 XX
 AC ABB78160;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Amino acid sequence of a YggX homologue.
 XX
 KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 XX
 OS Unidentified.
 XX

PN US2002072118-A1.
 XX 13-JUN-2002.
 PD
 XX
 PF 18-SEP-2001; 2001US-00955502.
 XX
 XX 22-SEP-2000; 2000US-0234588P.
 PR
 XX (DOWN/) DOWNS D.
 XX (GRAL/) GRALNICK J A.
 PA
 PA
 XX Downs D, Gralnick JA;
 XX WPI; 2002-589476/63.
 XX
 DR
 XX
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 PT
 XX Example; Fig 1A; 16pp; English.
 PS
 XX The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 88 AA;
 SQ
 Query Match 34.5%; Score 48; DB 5; Length 88;
 Best Local Similarity 23.6%; Pred. No. 0.0039;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 Qy 1 MRRXXXCXNXXXXXXXXXXPPXXXXXGXXXXXXXXXXXXXXXXXXXXXQTLXNEXXL 55
 Db 1 MSRTIFCTYLRDAGQDFQLYPGELGKRIYNEISKDAWAQWQHKTMLINEKKL 55
 RESULT 5
 ABB78161
 ID ABB78161 standard; protein; 91 AA.
 XX
 AC ABB78161;
 XX
 XX 05-NOV-2002 (first entry)
 DT
 XX Amino acid sequence of a YggX homologue.
 DE
 XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 KW
 XX Unidentified.
 OS
 XX US2002072118-A1.
 PN
 XX 13-JUN-2002.
 PD
 XX 18-SEP-2001; 2001US-00955502.
 PF
 XX 22-SEP-2000; 2000US-0234588P.
 PR
 XX (DOWN/) DOWNS D.
 XX (GRAL/) GRALNICK J A.
 PA
 PA Downs D, Gralnick JA;
 XX WPI; 2002-589476/63.
 XX
 DR
 XX
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 PT
 XX Example; Fig 1A; 16pp; English.
 PS
 XX The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 88 AA;
 SQ
 Query Match 34.5%; Score 48; DB 5; Length 88;
 Best Local Similarity 23.6%; Pred. No. 0.0039;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 Qy 1 MRRXXXCXNXXXXXXXXXXPPXXXXXGXXXXXXXXXXXXXXXXXXXXXQTLXNEXXL 55
 Db 1 MSRTIFCTYLRDAGQDFQLYPGELGKRIYNEISKDAWAQWQHKTMLINEKKL 55
 RESULT 5
 ABB78161
 ID ABB78161 standard; protein; 91 AA.
 XX
 AC ABB78161;
 XX
 XX 05-NOV-2002 (first entry)
 DT
 XX Amino acid sequence of a YggX homologue.
 DE
 XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 KW
 XX Unidentified.
 OS
 XX US2002072118-A1.
 PN
 XX 13-JUN-2002.
 PD
 XX 18-SEP-2001; 2001US-00955502.
 PF
 XX 22-SEP-2000; 2000US-0234588P.
 PR
 XX (DOWN/) DOWNS D.
 XX (GRAL/) GRALNICK J A.
 PA
 PA Downs D, Gralnick JA;
 XX WPI; 2002-589476/63.
 XX
 DR
 XX
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 PT
 XX Example; Fig 1A; 16pp; English.
 PS
 XX The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 91 AA;
 SQ

PT Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 XX
 XX Example; Fig 1A; 16pp; English.
 PS
 XX The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 91 AA;
 SQ
 Query Match 34.5%; Score 48; DB 5; Length 91;
 Best Local Similarity 23.6%; Pred. No. 0.004; 42; Indels 0; Gaps 0;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 Qy 1 MRRXXXCXNXXXXXXXXXXPPXXXXXGXXXXXXXXXXXXXXXXXXXXXQTLXNEXXL 55
 Db 1 MSRTIFCTYLRDAGQDFQLYPGELGKRIYNEISKDAWAQWQHKTMLINEKKL 55
 RESULT 6
 ABB78158
 ID ABB78158 standard; protein; 91 AA.
 XX
 AC ABB78158;
 XX
 XX 05-NOV-2002 (first entry)
 DT
 XX Amino acid sequence of a YggX homologue.
 DE
 XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 KW
 XX Unidentified.
 OS
 XX US2002072118-A1.
 PN
 XX 13-JUN-2002.
 PD
 XX 18-SEP-2001; 2001US-00955502.
 PF
 XX 22-SEP-2000; 2000US-0234588P.
 PR
 XX (DOWN/) DOWNS D.
 XX (GRAL/) GRALNICK J A.
 PA
 PA Downs D, Gralnick JA;
 XX WPI; 2002-589476/63.
 XX
 DR
 XX
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 PT
 XX Example; Fig 1A; 16pp; English.
 PS
 XX The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 91 AA;
 SQ

PD 13-JUN-2002.
 XX
 PF 18-SEP-2001; 2001US-00955502.
 XX
 PR 22-SEP-2000; 2000US-0234588P.
 XX
 PA (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 PI Downs D, Gralnick JA;
 XX
 XX WPI; 2002-589476/63.
 DR
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 XX
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 91 AA;
 SQ

Query Match 34.5%; Score 48; DB 5; Length 91;
 Best Local Similarity 23.6%; Pred. No. 0.004;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 Qy 1 MRRXXCXXXXXXXPPXXGXXXXXXXGXXXXXXXQTXLNEXXL 55
 Db 1 MSRTIFCYLQREAGQDFQLYPGLGKRIYNEISKAWAQWQHQTMLINEKKL 55

RESULT 10
 ABB78157
 ID ABB78157 standard; protein; 91 AA.
 XX
 AC ABB78157;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Amino acid sequence of a YggX homologue.
 XX
 KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 XX
 OS Unidentified.
 XX
 PN US2002072118-A1.
 XX
 PD 13-JUN-2002.
 XX
 PF 18-SEP-2001; 2001US-00955502.
 XX
 PR 22-SEP-2000; 2000US-0234588P.
 XX
 PA (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 PI Downs D, Gralnick JA;
 XX
 XX WPI; 2002-589476/63.
 DR
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 XX
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 91 AA;
 SQ

Query Match 34.5%; Score 48; DB 5; Length 91;
 Best Local Similarity 23.6%; Pred. No. 0.004;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 Qy 1 MRRXXCXXXXXXXPPXXGXXXXXXXGXXXXXXXQTXLNEXXL 55
 Db 1 MSRTIFCYLQREAGQDFQLYPGLGKRIYNEISKAWAQWQHQTMLINEKKL 55

RESULT 10
 ABB78157
 ID ABB78157 standard; protein; 91 AA.
 XX
 AC ABB78157;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Amino acid sequence of a YggX homologue.
 XX
 KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 XX
 OS Unidentified.
 XX
 PN US2002072118-A1.
 XX
 PD 13-JUN-2002.
 XX
 PF 18-SEP-2001; 2001US-00955502.
 XX
 PR 22-SEP-2000; 2000US-0234588P.
 XX
 PA (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 PI Downs D, Gralnick JA;
 XX
 XX WPI; 2002-589476/63.
 DR
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 XX
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 91 AA;
 SQ

PT protein identified from Salmonella enterica Serovar Typhimurium.
 XX
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 91 AA;
 SQ

Query Match 34.5%; Score 48; DB 5; Length 91;
 Best Local Similarity 23.6%; Pred. No. 0.004;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 Qy 1 MRRXXCXXXXXXXPPXXGXXXXXXXGXXXXXXXQTXLNEXXL 55
 Db 1 MSRTIFCYLQREAGQDFQLYPGLGKRIYNEISKAWAQWQHQTMLINEKKL 55

RESULT 11
 ABB78156
 ID ABB78156 standard; protein; 91 AA.
 XX
 AC ABB78156;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Amino acid sequence of a YggX homologue.
 XX
 KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 XX
 OS Escherichia coli.
 XX
 PN US2002072118-A1.
 XX
 PD 13-JUN-2002.
 XX
 PF 18-SEP-2001; 2001US-00955502.
 XX
 PR 22-SEP-2000; 2000US-0234588P.
 XX
 PA (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 PI Downs D, Gralnick JA;
 XX
 XX WPI; 2002-589476/63.
 DR
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 XX
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 91 AA;
 SQ

```

SQ Sequence 91 AA;
Query Match 34.5%; Score 48; DB 5; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.004;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MPRXXXCXKXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db 1 MSRTIFCTFLQREAGDQFQLYFGELGKRIYNEISKEAWAQHQTMLVNEKKL 55

RESULT 12
ABB78151
ID ABB78151 standard; protein; 87 AA.
XX
AC ABB78151;
XX
DT 05-NOV-2002 (first entry)
XX
DE Amino acid sequence of a YggX homologue.
XX
KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
KW hydroxyl radical; DNA damage; YggX homologue.
XX
OS Haemophilus influenzae.
XX
PN US2002072118-A1.
XX
PD 13-JUN-2002.
XX
PF 18-SEP-2001; 2001US-00955502.
XX
PR 22-SEP-2000; 2000US-0234588P.
XX
PA (DOWN/) DOWNS D.
PA (GRAL/) GRALNICK J A.
XX
PI Downs D, Gralnick JA;
XX
PI WPI; 2002-589476/63.
XX
PT Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
PT cell, comprises engineering the cell to produce more YggX protein, a
PT protein identified from Salmonella enterica serovar Typhimurium.
XX
PS Example; Fig 1A; 16pp; English.
XX
CC The specification describes a method for reducing superoxide damage to a
CC cell. The method comprises engineering the cell to produce more than the
CC native amount of YggX protein (a protein identified from Salmonella
CC enterica serovar typhimurium) or its homolog, where the cells are
CC rendered more resistant to superoxide damage. YggX reduces the oxidation
CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
CC clusters. The resulting decrease in free-iron levels generates fewer
CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
CC cell. ABB78147-78 represent YggX homologues
XX
PS Sequence 87 AA;
Query Match 33.8%; Score 47; DB 5; Length 87;
Best Local Similarity 23.6%; Pred. No. 0.0065;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MPRXXXCXKXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db 1 MARIVFCEYLKKEAGLDQFQLYFGELGKRIFDSVSKQAWGEWIKQTMLVNEKKL 55

RESULT 14
ABB78153
ID ABB78153 standard; protein; 87 AA.
XX
AC ABB78153;
XX
DT 05-NOV-2002 (first entry)
XX
DE Amino acid sequence of a YggX homologue.
XX
KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
KW hydroxyl radical; DNA damage; YggX homologue.
XX
OS Unidentified.
XX
PN US2002072118-A1.
XX
PD 13-JUN-2002.
XX

```


PF 18-SEP-2001; 2001US-00955502.
XX
PR 22-SEP-2000; 2000US-0234588P.
XX
PA (DOWN/) DOWNS D.
PA (GRAL/) GRALNICK J A.
XX
XX
PI Downs D, Gralnick JA;
XX
XX WPI; 2002-589476/63.
XX
XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
PT cell, comprises engineering the cell to produce more YggX protein, a
PT protein identified from Salmonella enterica Serovar Typhimurium.
XX
XX Example; Fig 1A; 16pp; English.
XX
XX The specification describes a method for reducing superoxide damage to a
CC cell. The method comprises engineering the cell to produce more than the
CC native amount of YggX protein (a protein identified from Salmonella
CC enterica serovar typhimurium) or its homolog, where the cells are
CC rendered more resistant to superoxide damage. YggX reduces the oxidation
CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
CC clusters. The resulting decrease in free-iron levels generates fewer
CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
CC cell. ABB78147-78 represent YggX homologues
XX
SQ Sequence 87 AA;

Query Match 33.8%; Score 47; DB 5; Length 87;
Best Local Similarity 23.6%; Pred. No. 0.0065;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXXCXNXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXQPLXNEXXL 55
Db 1 MARVVFCEYLKKEAGLDLFQLYPGELGKRFNSISKQAEWIKKQTMVNEKKL 55

RESULT 15
ABB78155
ID ABB78155 standard; protein; 90 AA.
XX
AC ABB78155;
XX
DT 05-NOV-2002 (first entry)
XX
DE Amino acid sequence of a YggX homologue.
XX
KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
KW hydroxyl radical; DNA damage; YggX homologue.
XX
OS Vibrio cholerae.
XX
XX US2002072118-A1.
XX
PN 13-JUN-2002.
XX
PF 18-SEP-2001; 2001US-00955502.
XX
PR 22-SEP-2000; 2000US-0234588P.
XX
PA (DOWN/) DOWNS D.
PA (GRAL/) GRALNICK J A.
XX
XX Downs D, Gralnick JA;
XX
XX WPI; 2002-589476/63.
XX
XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
PT cell, comprises engineering the cell to produce more YggX protein, a
PT protein identified from Salmonella enterica Serovar Typhimurium.
XX

PS Example; Fig 1A; 16pp; English.
XX
CC The specification describes a method for reducing superoxide damage to a
CC cell. The method comprises engineering the cell to produce more than the
CC native amount of YggX protein (a protein identified from Salmonella
CC enterica serovar typhimurium) or its homolog, where the cells are
CC rendered more resistant to superoxide damage. YggX reduces the oxidation
CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
CC clusters. The resulting decrease in free-iron levels generates fewer
CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
CC cell. ABB78147-78 represent YggX homologues
XX
SQ Sequence 90 AA;

Query Match 33.8%; Score 47; DB 5; Length 90;
Best Local Similarity 23.6%; Pred. No. 0.0067;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXXCXNXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXQPLXNEXXL 55
Db 1 MARTVFCTRLQKEADGLDFQLYPGELGKRFIDNICKEAWAQWTKQTMLINEKKL 55

RESULT 16
ABB78166
ID ABB78166 standard; protein; 76 AA.
XX
AC ABB78166;
XX
DT 05-NOV-2002 (first entry)
XX
DE Amino acid sequence of a YggX homologue.
XX
KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
KW hydroxyl radical; DNA damage; YggX homologue.
XX
OS Unidentified.
XX
XX US2002072118-A1.
XX
PN 13-JUN-2002.
XX
PF 18-SEP-2001; 2001US-00955502.
XX
PR 22-SEP-2000; 2000US-0234588P.
XX
XX (DOWN/) DOWNS D.
XX (GRAL/) GRALNICK J A.
XX
XX Downs D, Gralnick JA;
XX
XX WPI; 2002-589476/63.
XX
XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
PT cell, comprises engineering the cell to produce more YggX protein, a
PT protein identified from Salmonella enterica Serovar Typhimurium.
XX
XX Example; Fig 1A; 16pp; English.
XX
XX The specification describes a method for reducing superoxide damage to a
CC cell. The method comprises engineering the cell to produce more than the
CC native amount of YggX protein (a protein identified from Salmonella
CC enterica serovar typhimurium) or its homolog, where the cells are
CC rendered more resistant to superoxide damage. YggX reduces the oxidation
CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
CC clusters. The resulting decrease in free-iron levels generates fewer
CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
CC cell. ABB78147-78 represent YggX homologues
XX
SQ Sequence 76 AA;

PF 18-SEP-2001; 2001US-009555502.
XX
XX
PR 22-SEP-2000; 2000US-0234588P.
XX
XX
PA (DOWN/) DOWNS D.
XX (GRAL/) GRALNICK J A.
XX
XX
PI Downs D, Gralnick JA;
XX
XX WPI; 2002-589476/63.
XX
XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
PT cell, comprises engineering the cell to produce more YggX protein, a
PT protein identified from Salmonella enterica Serovar Typhimurium.
XX
XX Example; Fig 1A; 16pp; English.
XX
XX The specification describes a method for reducing superoxide damage to a
CC cell. The method comprises engineering the cell to produce more than the
CC native amount of YggX protein (a protein identified from Salmonella
CC enterica serovar typhimurium) or its homolog, where the cells are
CC rendered more resistant to superoxide damage. YggX reduces the oxidation
CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
CC clusters. The resulting decrease in free-iron levels generates fewer
CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
CC cell. ABB78147-78 represent YggX homologues
XX
XX Sequence 88 AA;
SQ
Query Match 33.1%; Score 46; DB 5; Length 88;
Best Local Similarity 23.6%; Pred. No. 0.011;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 MRRXXCCXXXXXXXPPXGXXXXXXXQXXXXXXXXXQXXLXNEXXL 55
Db 1 MARTVNCVHLNKEADGLDFQLYPGDLGKRIFDNISKWAGLWQXKQTMLINEKKL 55
RESULT 20
AB78173
ID ABB78173 standard; protein; 88 AA.
XX
XX ABB78173;
XX
XX 05-NOV-2002 (first entry)
XX
XX Amino acid sequence of a YggX homologue.
XX
XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
KW hydroxyl radical; DNA damage; YggX homologue.
XX
XX Neisseria meningitidis.
OS
XX US2002072118-A1.
XX
XX 13-JUN-2002.
XX
XX 18-SEP-2001; 2001US-009555502.
XX
XX 22-SEP-2000; 2000US-0234588P.
XX
XX (DOWN/) DOWNS D.
XX (GRAL/) GRALNICK J A.
XX
XX Downs D, Gralnick JA;
XX
XX WPI; 2002-589476/63.
XX
XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
PT cell, comprises engineering the cell to produce more YggX protein, a
PT protein identified from Salmonella enterica Serovar Typhimurium.
XX

PS Example; Fig 1A; 16pp; English.
XX
XX The specification describes a method for reducing superoxide damage to a
CC cell. The method comprises engineering the cell to produce more than the
CC native amount of YggX protein (a protein identified from Salmonella
CC enterica serovar typhimurium) or its homolog, where the cells are
CC rendered more resistant to superoxide damage. YggX reduces the oxidation
CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
CC clusters. The resulting decrease in free-iron levels generates fewer
CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
CC cell. ABB78147-78 represent YggX homologues
XX
XX Sequence 88 AA;
SQ
Query Match 33.1%; Score 46; DB 5; Length 88;
Best Local Similarity 30.3%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
Qy 23 PXXGXXXXXXXQXXXXXXXXXQXXLXNEXXL 55
Db 23 PNELGKRIFFENVSQBAWAAWTRHQTMLINENEL 55
RESULT 21
ABP77219
ID ABP77219 standard; protein; 88 AA.
XX
XX ABP77219;
XX
XX 07-MAR-2003 (first entry)
XX
XX N. gonorrhoeae amino acid sequence SEQ ID 968.
XX
XX Antibacterial; infection; vaccine; gene therapy.
XX
XX Neisseria gonorrhoeae.
XX
XX WO200279243-A2.
XX
XX 10-OCT-2002.
XX
XX 12-FEB-2002; 2002WO-IB002069.
XX
XX 12-FEB-2001; 2001GB-00003424.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Fontana MR, Pizza M, Massignani V, Monaci E;
XX
XX WPI; 2003-058415/05.
XX
XX N-PSDB; AB238189.
XX
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoea infection.
XX
XX Disclosure; Page 258; 815pp; English.
XX
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention
XX
XX Sequence 88 AA;
SQ
Query Match 33.1%; Score 46; DB 6; Length 88;
Best Local Similarity 30.3%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 23 PXXXXGXXXXXXXWXXWXXQTXLXNEXXL 55
 Db 23 PNELGKRFENVSQEAAWTRHQTMLINERL 55

RESULT 22

ABBT78168
 ID ABB78168 standard; protein; 90 AA.

XX AC ABB78168;
 XX 05-NOV-2002 (first entry)

XX Amino acid sequence of a YggX homologue.

XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 XX hydroxyl radical; DNA damage; YggX homologue.

XX Pseudomonas syringae.

XX US2002072118-A1.

XX 13-JUN-2002.

XX 18-SEP-2001; 2001US-00955502.

XX 22-SEP-2000; 2000US-0234588P.

XX (DOWN/) DOWNS D.

XX (GRAL/) GRALNICK J A.

XX Downs D, Gralnick JA;

XX WPI; 2002-589476/63.

XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 cell, comprises engineering the cell to produce more YggX protein, a
 protein identified from Salmonella enterica Serovar Typhimurium.

XX Example; Fig 1A; 16pp; English.

XX The specification describes a method for reducing superoxide damage to a
 cell. The method comprises engineering the cell to produce more than the
 native amount of YggX protein (a protein identified from Salmonella
 enterica serovar typhimurium) or its homolog, where the cells are
 rendered more resistant to superoxide damage. YggX reduces the oxidation
 of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 clusters. The resulting decrease in free-iron levels generates fewer
 hydroxyl radicals and thus reduced DNA damage. The method is useful for
 reducing superoxide damage in a bacterial, yeast, mammalian or plant
 cell. ABB78147-78 represent YggX homologues

XX Sequence 90 AA;

Query Match 33.1%; Score 46; DB 5; Length 90;
 Best Local Similarity 23.6%; Pred. No. 0.011;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MXXRXXCXXXXXXXFPXXGXXXXXXXWXXWXXQTXLXNEXXL 55

Db 1 MTRTVMCRVKYKELPGLERAPYFGAKGEDI FNHVSQKAWADWQHQTLLINERL 55

RESULT 23

ABBT78150
 ID ABB78150 standard; protein; 91 AA.

XX ABB78150;

XX 05-NOV-2002 (first entry)

XX Amino acid sequence of a YggX homologue.

KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.

XX Unidentified.

XX US2002072118-A1.

XX 13-JUN-2002.

XX 18-SEP-2001; 2001US-00955502.

XX 22-SEP-2000; 2000US-0234588P.

XX (DOWN/) DOWNS D.

XX (GRAL/) GRALNICK J A.

XX Downs D, Gralnick JA;

XX WPI; 2002-589476/63.

XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 cell, comprises engineering the cell to produce more YggX protein, a
 protein identified from Salmonella enterica Serovar Typhimurium.

XX Example; Fig 1A; 16pp; English.

XX The specification describes a method for reducing superoxide damage to a
 cell. The method comprises engineering the cell to produce more than the
 native amount of YggX protein (a protein identified from Salmonella
 enterica serovar typhimurium) or its homolog, where the cells are
 rendered more resistant to superoxide damage. YggX reduces the oxidation
 of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 clusters. The resulting decrease in free-iron levels generates fewer
 hydroxyl radicals and thus reduced DNA damage. The method is useful for
 reducing superoxide damage in a bacterial, yeast, mammalian or plant
 cell. ABB78147-78 represent YggX homologues

XX Sequence 91 AA;

Query Match 33.1%; Score 46; DB 5; Length 91;
 Best Local Similarity 23.6%; Pred. No. 0.011;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MXXRXXCXXXXXXXFPXXGXXXXXXXWXXWXXQTXLXNEXXL 55

Db 1 MARWVFCERLQBAEGLDFQLYPGELGKRFIDSIKQAWGEMWKKQTMVLNKKL 55

RESULT 24

ADA34169
 ID ADA34169 standard; protein; 92 AA.

XX ADA34169;

XX 20-NOV-2003 (first entry)

XX Acinetobacter baumannii protein #1330.

XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 plant biocontrol agent.

XX Acinetobacter baumannii.

XX US6562958-B1.

XX 13-MAY-2003.

XX 04-JUN-1999; 99US-00328352.

XX 09-JUN-1998; 98US-0088701P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Breton G, Bush D;
 XX WPI; 2003-576092/54.
 DR N-PSDB; ADA30043.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 XX Example; SEQ ID NO 5456; 328pp; English.
 PS
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX
 XX Sequence 92 AA;
 SQ
 Query Match 33.1%; Score 46; DB 6; Length 92;
 Best Local Similarity 23.6%; Pred. No. 0.012;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 QY 1 MRRXXXCXGXXXXXXXPPXXGXXXXXXXGXXXXXXXQTXLXNEXXL 55
 Db 4 MSRVFCRKQXKMEGLDFAPFGAKGQEPFENVSKQAWQEWLQHTTLINERKL 58
 RESULT 25
 ABB78175
 ID ABB78175 standard; protein; 87 AA.
 XX
 AC ABB78175;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Amino acid sequence of a YggX homologue.
 XX
 KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 XX
 OS Unidentified.
 XX
 XX US2002072118-A1.
 XX
 PD 13-JUN-2002.
 XX
 PF 18-SEP-2001; 2001US-00955502.
 XX
 PR 22-SEP-2000; 2000US-0234588P.
 XX
 PA (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 PI Downs D, Gralnick JA;
 XX
 DR WPI; 2002-589476/63.
 XX
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 XX
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homologue, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 87 AA;
 SQ
 Query Match 32.4%; Score 45; DB 5; Length 87;
 Best Local Similarity 30.3%; Pred. No. 0.018;
 Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 23 PXXGXXXXXXXGXXXXXXXQTXLXNEXXL 55
 Db 1 MRRTHCAKLGKAEGLDFPPLPGELGKRLYESVSKQAWQDWLQKQTWLINENRL 55

CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 87 AA;
 SQ
 Query Match 32.4%; Score 45; DB 5; Length 87;
 Best Local Similarity 23.6%; Pred. No. 0.018; 42; Indels 0; Gaps 0;
 Matches 13; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 1 MRRXXXCXGXXXXXXXPPXXGXXXXXXXGXXXXXXXQTXLXNEXXL 55
 Db 1 MRRTHCAKLGKAEGLDFPPLPGELGKRLYESVSKQAWQDWLQKQTWLINENRL 55
 RESULT 26
 ABB78169
 ID ABB78169 standard; protein; 87 AA.
 XX
 AC ABB78169;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Amino acid sequence of a YggX homologue.
 XX
 KW Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 XX
 OS Pseudomonas putida.
 XX
 XX US2002072118-A1.
 XX
 PD 13-JUN-2002.
 XX
 PF 18-SEP-2001; 2001US-00955502.
 XX
 PR 22-SEP-2000; 2000US-0234588P.
 XX
 PA (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 PI Downs D, Gralnick JA;
 XX
 DR WPI; 2002-589476/63.
 XX
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 PT cell, comprises engineering the cell to produce more YggX protein, a
 PT protein identified from Salmonella enterica Serovar Typhimurium.
 XX
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homologue, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 87 AA;
 SQ
 Query Match 32.4%; Score 45; DB 5; Length 87;
 Best Local Similarity 30.3%; Pred. No. 0.018;
 Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 23 PXXGXXXXXXXGXXXXXXXQTXLXNEXXL 55

Db 23 PGAKGQDIFEHISQKAWADWQHQTMLINEKRL 55

RESULT 27

ABB78177
ID ABB78177 standard; protein; 87 AA.

AC ABB78177;

XX 05-NOV-2002 (first entry)

XX Amino acid sequence of a YggX homologue.

XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
KW hydroxyl radical; DNA damage; YggX homologue.

XX Methylococcus capsulatus.

OS US2002072118-A1.

XX 13-JUN-2002.

XX 18-SEP-2001; 2001US-00955502.

XX 22-SEP-2000; 2000US-0234588P.

XX (DOWN/) DOWNS D.

PA (GRAL/) GRALNICK J A.

PI Downs D, Gralnick JA;

XX WPI; 2002-589476/63.

XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
PT cell, comprises engineering the cell to produce more YggX protein, a
PT protein identified from Salmonella enterica Serovar Typhimurium.

XX Example; Fig 1A; 16pp; English.

XX The specification describes a method for reducing superoxide damage to a
CC cell. The method comprises engineering the cell to produce more than the
CC native amount of YggX protein (a protein identified from Salmonella
CC enterica serovar typhimurium) or its homolog, where the cells are
CC rendered more resistant to superoxide damage. YggX reduces the oxidation
CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
CC clusters. The resulting decrease in free-iron levels generates fewer
CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
CC cell. ABB78147-78 represent YggX homologues.

XX Sequence 87 AA;

Query Match

Best Local Similarity 32.4%; Score 45; DB 5; Length 87;
Matches 14; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 MRRXXXXXXXXXXXXXXFPXXXXXXXQXXXXXXXXXXXXXXQTXLXNEXLXXXX 60

Db 1 MARRIICAKLGIEADGLDAPPFGPGQGRIFEHVSKEAWQDLKLQDTMLINEHRLTPPEA 60

QY 61 XXR 63

Db 61 SAR 63

RESULT 28

ABB78176
ID ABB78176 standard; protein; 87 AA.

AC ABB78176;

XX 29-AUG-2003 (revised)

DT 05-NOV-2002 (first entry)

XX

DE Amino acid sequence of a YggX homologue.

XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
KW hydroxyl radical; DNA damage; YggX homologue.

OS Acidithiobacillus ferrooxidans.

XX US2002072118-A1.

XX 13-JUN-2002.

XX 18-SEP-2001; 2001US-00955502.

XX 22-SEP-2000; 2000US-0234588P.

XX (DOWN/) DOWNS D.

PA (GRAL/) GRALNICK J A.

PI Downs D, Gralnick JA;

XX WPI; 2002-589476/63.

XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
PT cell, comprises engineering the cell to produce more YggX protein, a
PT protein identified from Salmonella enterica Serovar Typhimurium.

XX Example; Fig 1A; 16pp; English.

XX The specification describes a method for reducing superoxide damage to a
CC cell. The method comprises engineering the cell to produce more than the
CC native amount of YggX protein (a protein identified from Salmonella
CC enterica serovar typhimurium) or its homolog, where the cells are
CC rendered more resistant to superoxide damage. YggX reduces the oxidation
CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
CC clusters. The resulting decrease in free-iron levels generates fewer
CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
CC cell. ABB78147-78 represent YggX homologues. (Updated on 29-AUG-2003 to
CC standardise OS field)

XX Sequence 87 AA;

Query Match

Best Local Similarity 32.4%; Score 45; DB 5; Length 87;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 23 PXXXXXXXXXXXXXXXXXQXXXXXXXXXXXXXXQTXLXNEXL 55

Db 23 PGALGARIYQEVSVKEAWQGLKHQTMLINEYRL 55

RESULT 29

ABB78149

ID ABB78149 standard; protein; 86 AA.

XX ABB78149;

XX 05-NOV-2002 (first entry)

XX Amino acid sequence of a YggX homologue.

XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
KW hydroxyl radical; DNA damage; YggX homologue.

XX Bordetella bronchiseptica.

XX US2002072118-A1.

XX 13-JUN-2002.

XX 18-SEP-2001; 2001US-00955502.

XX

PR 22-SEP-2000; 2000US-0234588P.
 XX (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 XX Downs D, Gralnick JA;
 PI WPI; 2002-589476/63.
 XX
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 XX cell, comprises engineering the cell to produce more YggX protein, a
 XX protein identified from Salmonella enterica Serovar Typhimurium.
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 86 AA;
 SQ
 Query Match 31.7%; Score 44; DB 5; Length 86;
 Best Local Similarity 26.8%; Pred. No. 0.031;
 Matches 11; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
 QY 23 PXXGXGXXXXXXXWXXWXXQTXLXNEXLXXXXXXXR 63
 DB 23 FGELGTRIQWQISKEAWEEWQIQTLVNLNADARAR 63
 RESULT 30
 ABB78148
 ID ABB78148 standard; protein; 87 AA.
 XX
 XX ABB78148;
 AC
 XX
 DT 05-NOV-2002 (first entry)
 XX
 XX Amino acid sequence of a YggX homologue.
 DE
 XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 XX
 XX Bordetella parapertussis.
 OS
 XX US2002072118-A1.
 PN
 XX 13-JUN-2002.
 PD
 XX 18-SEP-2001; 2001US-00955502.
 PF
 XX 22-SEP-2000; 2000US-0234588P.
 PR
 XX (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 XX Downs D, Gralnick JA;
 PI WPI; 2002-589476/63.
 XX
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 XX cell, comprises engineering the cell to produce more YggX protein, a
 XX protein identified from Salmonella enterica Serovar Typhimurium.
 PS Example; Fig 1A; 16pp; English.
 XX

CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 87 AA;
 SQ
 Query Match 31.7%; Score 44; DB 5; Length 87;
 Best Local Similarity 26.8%; Pred. No. 0.031;
 Matches 11; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
 QY 23 PXXGXGXXXXXXXWXXWXXQTXLXNEXLXXXXXXXR 63
 DB 23 FGELGTRIQWQISKEAWEEWQIQTLVNLNADARAR 63
 RESULT 31
 ABB78170
 ID ABB78170 standard; protein; 87 AA.
 XX
 XX ABB78170;
 AC
 XX
 DT 05-NOV-2002 (first entry)
 XX
 XX Amino acid sequence of a YggX homologue.
 DE
 XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
 KW hydroxyl radical; DNA damage; YggX homologue.
 XX
 XX Unidentified.
 OS
 XX US2002072118-A1.
 PN
 XX 13-JUN-2002.
 PD
 XX 18-SEP-2001; 2001US-00955502.
 PF
 XX 22-SEP-2000; 2000US-0234588P.
 PR
 XX (DOWN/) DOWNS D.
 PA (GRAL/) GRALNICK J A.
 XX
 XX Downs D, Gralnick JA;
 PI WPI; 2002-589476/63.
 XX
 XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
 XX cell, comprises engineering the cell to produce more YggX protein, a
 XX protein identified from Salmonella enterica Serovar Typhimurium.
 PS Example; Fig 1A; 16pp; English.
 XX
 CC The specification describes a method for reducing superoxide damage to a
 CC cell. The method comprises engineering the cell to produce more than the
 CC native amount of YggX protein (a protein identified from Salmonella
 CC enterica serovar typhimurium) or its homolog, where the cells are
 CC rendered more resistant to superoxide damage. YggX reduces the oxidation
 CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
 CC clusters. The resulting decrease in free-iron levels generates fewer
 CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
 CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
 CC cell. ABB78147-78 represent YggX homologues
 XX
 XX Sequence 87 AA;
 SQ
 Query Match 31.7%; Score 44; DB 5; Length 87;
 Best Local Similarity 30.3%; Pred. No. 0.031;

PN US2002072118-A1.
XX 13-JUN-2002.
XX 18-SEP-2001; 2001US-009555502.
XX 22-SEP-2000; 2000US-0234588P.
XX (DOWN/) DOWNS D.
XX (GRAL/) GRALNICK J A.
XX Downs D, Gralnick JA;
XX WPI; 2002-589476/63.
XX Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
PT cell, comprises engineering the cell to produce more YggX protein, a
PT protein identified from Salmonella enterica Serovar Typhimurium.
XX Claim 8; Page 7; 16pp; English.
XX The specification describes a method for reducing superoxide damage to a
CC cell. The method comprises engineering the cell to produce more than the
CC native amount of YggX protein (a protein identified from Salmonella
CC enterica serovar typhimurium) or its homolog, where the cells are
CC rendered more resistant to superoxide damage. YggX reduces the oxidation
CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
CC clusters. The resulting decrease in free-iron levels generates fewer
CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
CC cell. The present sequence represents a YggX homologue consensus sequence
XX Sequence 62 AA;
Query Match 30.2%; Score 42; DB 5; Length 62;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 39 WXXWXXXQTXLNEXXL 55
Db 36 WXXWXXXQTXLNEXXL 52
RESULT 35
ABB78174
ID ABB78174 standard; protein; 87 AA.
XX AC ABB78174;
XX 05-NOV-2002 (first entry)
XX Amino acid sequence of a YggX homologue.
XX Superoxide damage; cell; YggX; Salmonella enterica serovar typhimurium;
KW hydroxyl radical; DNA damage; YggX homologue.
XX Unidentified.
XX US2002072118-A1.
XX 13-JUN-2002.
XX 18-SEP-2001; 2001US-009555502.
XX 22-SEP-2000; 2000US-0234588P.
XX (DOWN/) DOWNS D.
XX (GRAL/) GRALNICK J A.
XX Downs D, Gralnick JA;
XX WPI; 2002-589476/63.
XX

PT Reducing superoxide damage to a cell, e.g. bacterial, mammalian or plant
PT cell, comprises engineering the cell to produce more YggX protein, a
PT protein identified from Salmonella enterica Serovar Typhimurium.
XX Example; Fig 1A; 16pp; English.
XX The specification describes a method for reducing superoxide damage to a
CC cell. The method comprises engineering the cell to produce more than the
CC native amount of YggX protein (a protein identified from Salmonella
CC enterica serovar typhimurium) or its homolog, where the cells are
CC rendered more resistant to superoxide damage. YggX reduces the oxidation
CC of (Fe-S) clusters, and thus reduces the loss of Fe(II) ions from
CC clusters. The resulting decrease in free-iron levels generates fewer
CC hydroxyl radicals and thus reduced DNA damage. The method is useful for
CC reducing superoxide damage in a bacterial, yeast, mammalian or plant
CC cell. ABB78147-78 represent YggX homologues
XX Sequence 87 AA;
Query Match 28.1%; Score 39; DB 5; Length 87;
Best Local Similarity 21.8%; Pred. No. 0.43;
Matches 12; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
Qy 1 MXXRXXCXXXXXXXPPPPGSLGKRLYESVSKQAWQDWLKKQIMLINENRL 55
Db 1 MARMTHCAKLGKXEBGLDPPPLGSLGKRLYESVSKQAWQDWLKKQIMLINENRL 55
RESULT 36
ABB70457
ID ABB70457 standard; protein; 1767 AA.
XX AC ABB70457;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 38163.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL14560.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX Disclosure; SEQ ID NO 38163; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1767 AA;

Query Match 25.2%; Score 35; DB 4; Length 1767;
 Best Local Similarity 35.7%; Pred. No. 56;
 Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
 Db 1183 WSRWSESYLRNE 1196

RESULT 37
 AAR10681
 ID AAR10681 standard; protein; 589 AA.
 XX
 AC AAR10681;
 XX
 XX 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 17-APR-1991 (first entry)
 XX
 XX Polyhydroxybutyrate polymerase enzyme.
 XX
 XX Polyester biopolymers; polyhydroxybutyrate; polyhydroxy alkanate;
 KW beta-ketothiolase; acetoacetyl CoA reductase.
 XX
 XX Ralstonia eutropha.
 OS
 XX
 PN WO9100917-A.
 XX
 PD 24-JAN-1991.
 XX
 PF 10-JUL-1989; 89US-00378155.
 XX
 PR 10-JUL-1989; 89US-00378155.
 XX
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 PA
 XX Peoples OP, Sinskey AJ;
 PI
 XX WPI; 1991-051341/07.
 DR N-PSDB; AAQ10502.
 XX
 XX Construction and modification of polyester bio;polymers - by introduction
 PT of poly-hydroxy-butyrate and -alkanoate genes into bacteria or plants.
 XX
 XX Disclosure; Fig 4; 64pp; English.

CC This Alcaligenes eutrophus polyhydroxybutyrate (PHB) polymerase enzyme is
 CC essential to the biosynthesis of PHB. The use of recombinant methods for
 CC producing such enzymes, required for polyester biopolymer synthesis,
 CC allows for the control and modification of the synthesis process. See
 CC also AAQ10499-501 and AAQ10503. (Updated on 25-MAR-2003 to correct PA
 CC field.) (Updated on 24-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 589 AA;

Query Match 24.5%; Score 34; DB 2; Length 589;
 Best Local Similarity 42.9%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
 Db 436 WYCWLRLRHTYLQNE 449

RESULT 38
 AAR32190
 ID AAR32190 standard; protein; 589 AA.
 XX

AC AAR32190;
 XX 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 30-MAY-1993 (first entry)
 XX
 DE Sequence encoded by the PHB synthase (phbC) gene of the
 DE polyhydroxybutyrate (PHB) operon.
 XX
 KW Operon; polyhydroxyalkanoate; polyhydroxybutyrate synthase.
 XX
 OS Ralstonia eutropha.
 XX
 PN WO9302187-A1.
 XX
 PD 04-FEB-1993.
 XX
 PF 13-JUL-1992; 92WO-US005786.
 XX
 PR 19-JUL-1991; 91US-00732243.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Somerville CR, Poirier Y, Dennis DE;
 XX
 DR WPI; 1993-058785/07.
 DR N-PSDB; AAQ36660.
 XX
 PT Transgenic plants producing poly:hydroxy-alkanoate polymer(s) - obt'd. by
 PT transformation with DNA encoding 3-ketothiolase, acetoacetyl-CoA
 PT reductase and PHA synthase.
 XX
 XX Disclosure; Fig 2; 70pp; English.

CC The nucleotide sequence of the PHB operon was obtained from Janes, B.
 CC Hollar, J. and Dennis, D. in Daves, E.A. (ed.) Novel Biodegradable
 CC Polymers, Kluwer Academic Publishers, 175-190 (1990). It contains the
 CC genes from PHB synthase, 3-ketothiolase and acetoacetyl-CoA reductase.
 CC The inventors claim a transgenic plant material contg. foreign DNA
 CC encoding a peptide which exhibits 3-ketothiolase activity, pref. where the
 CC DNA is an open reading from between nucleotides 2696-3877 (phb A gene),
 CC 842-2611 (phb C gene) or 3952-4692 (phb B gene) of the Alcaligenes
 CC eutrophus PHB operon. (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 24-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 589 AA;

Query Match 24.5%; Score 34; DB 2; Length 589;
 Best Local Similarity 42.9%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
 Db 436 WYCWLRLRHTYLQNE 449

RESULT 39
 AAE10892
 ID AAE10892 standard; protein; 589 AA.
 XX
 AC AAE10892;
 XX
 XX 11-SEP-2003 (revised)
 DT 18-DEC-2001 (first entry)
 XX
 DE Alcaligenes eutrophus polyhydroxybutyrate (PHB) synthase.
 XX
 KW Polyhydroxybutyrate synthase; PHB; sphingian; food product; dessert gel;
 KW jelly; jam; beverage; dairy product; gelling agent; rheological modifier;
 KW industrial application.
 XX
 OS Ralstonia eutropha.
 XX

PN W0200164897-A2.
XX 07-SEP-2001.
XX 02-MAR-2001; 2001WO-US007010.
XX 02-MAR-2000; 2000US-0186433P.
XX (KELC) CP KELCO US INC.
XX Bower S, Burke E, Harding N, Patel YN, Schneider JC, Meissner D;
PI Morrison N, Bezanson R;
XX WPI; 2001-589870/66.
XX Mutant strain of *Sphingomonas* species useful for producing sphingans, has
PT a mutation in the gene encoding a protein involved in polyhydroxybutyrate
PT synthesis.
XX
XX Example 1; Page 82-84; 98pp; English.
XX The invention relates to mutant strains of *Sphingomonas* species which
CC have a mutation in the gene encoding a protein involved in internal
CC storage polymer polyhydroxybutyrate (PHB) synthesis that allows the
CC mutant strains to produce PHB-deficient sphingans. Sphingans are capsular
CC polysaccharides secreted by bacteria of the genus *Sphingomonas*. The
CC invention also relates to a process for preparing clarified sphingans
CC which are useful as gelling agents in a variety of food applications for
CC improving the taste, texture, stability and appearance of food products
CC such as dessert gels, confectionery jellies, jams, dairy products,
CC beverages, films and coatings. The sphingans are also useful as
CC rheological modifier in industrial applications such as oil-field
CC drilling and cementitious systems. The present sequence is Alcaligenes
CC eutrophus polyhydroxybutyrate (PHB) synthase which is used for the
CC generation of *Sphingomonas* elodea phac fragment. phac DNA fragment
CC encodes polyhydroxybutyrate (PHB) synthase protein. (Updated on 11-SEP-
XX 2003 to standardise OS field)
XX Sequence 589 AA;
SQ

Query Match 24.5%; Score 34; DB 4; Length 589;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 436 WYCWYLRHTYLQNE 449

RESULT 40
ABU10991
ID ABU10991 standard; protein; 589 AA.
XX
XX AC ABU10991;
XX
XX 04-FEB-2003 (first entry)
XX Poly3-hydroxybutanoate synthase mutant associated protein #18.
XX
XX Poly3-hydroxyalkanoic acid; biodegradable polyester.
XX
XX Synthetic.
XX
XX JP2002199890-A.
XX
XX 16-JUL-2002.
XX
XX 28-FEB-2001; 2001JP-00054717.
XX
XX 23-OCT-2000; 2000JP-00322748.
XX
XX (RIKA) RIKAGAKU KENKYUSHO.
XX

Query Match 24.5%; Score 34; DB 4; Length 589;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 436 WYCWYLRHTYLQNE 449

RESULT 40
ABU10991
ID ABU10991 standard; protein; 589 AA.
XX
XX AC ABU10991;
XX
XX 04-FEB-2003 (first entry)
XX Poly3-hydroxybutanoate synthase mutant associated protein #18.
XX
XX Poly3-hydroxyalkanoic acid; biodegradable polyester.
XX
XX Synthetic.
XX
XX JP2002199890-A.
XX
XX 16-JUL-2002.
XX
XX 28-FEB-2001; 2001JP-00054717.
XX
XX 23-OCT-2000; 2000JP-00322748.
XX
XX (RIKA) RIKAGAKU KENKYUSHO.
XX

DR WPI; 2002-744015/81.
XX N-PSDB; ABX17307.
XX
XX Modification of a biodegradable polyester synthase, a mutant poly3-
PT hydroxybutanoate synthase, its preparation, a recombinant vector, a
PT transformant, preparation of a biodegradable ester polymer.
XX
XX Claim 13; Page 115-117; 124pp; Japanese.
XX
XX This invention relates to a novel method for the modification of an
CC enzyme participating to the biosynthesis of a poly3-hydroxyalkanoic acid
CC by modifying by recombinant DNA technology. The invention also comprises
CC a gene encoding the above mutant poly3-hydroxybutanoate synthase and a
CC recombinant vector containing the above gene. The method of the invention
CC may be used for the preparation of biodegradable polyesters. The present
CC sequence represents a protein used the method of the invention
XX
XX Sequence 589 AA;
SQ

Query Match 24.5%; Score 34; DB 5; Length 589;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 436 WYCWYLRHTYLQNE 449

RESULT 41
ABU10974
ID ABU10974 standard; protein; 589 AA.
XX
XX AC ABU10974;
XX
XX 04-FEB-2003 (first entry)
XX Poly3-hydroxybutanoate synthase mutant #1.
XX
XX Poly3-hydroxyalkanoic acid; biodegradable polyester; mutant; mutein.
XX
XX Synthetic.
XX
XX JP2002199890-A.
XX
XX 16-JUL-2002.
XX
XX 28-FEB-2001; 2001JP-00054717.
XX
XX 23-OCT-2000; 2000JP-00322748.
XX
XX (RIKA) RIKAGAKU KENKYUSHO.
XX
XX WPI; 2002-744015/81.
XX N-PSDB; ABX17290.
XX
XX Modification of a biodegradable polyester synthase, a mutant poly3-
PT hydroxybutanoate synthase, its preparation, a recombinant vector, a
PT transformant, preparation of a biodegradable ester polymer.
XX
XX Claim 13; Page 46-47; 124pp; Japanese.
XX
XX This invention relates to a novel method for the modification of an
CC enzyme participating to the biosynthesis of a poly3-hydroxyalkanoic acid
CC by modifying by recombinant DNA technology. The invention also comprises
CC a gene encoding the above mutant poly3-hydroxybutanoate synthase and a
CC recombinant vector containing the above gene. The method of the invention
CC may be used for the preparation of biodegradable polyesters. The present
CC sequence represents a protein used the method of the invention
XX
XX Sequence 589 AA;
SQ

Query Match 24.5%; Score 34; DB 5; Length 589;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 436 WYCWYLRHTYLQNE 449

RESULT 41
ABU10974
ID ABU10974 standard; protein; 589 AA.
XX
XX AC ABU10974;
XX
XX 04-FEB-2003 (first entry)
XX Poly3-hydroxybutanoate synthase mutant #1.
XX
XX Poly3-hydroxyalkanoic acid; biodegradable polyester; mutant; mutein.
XX
XX Synthetic.
XX
XX JP2002199890-A.
XX
XX 16-JUL-2002.
XX
XX 28-FEB-2001; 2001JP-00054717.
XX
XX 23-OCT-2000; 2000JP-00322748.
XX
XX (RIKA) RIKAGAKU KENKYUSHO.
XX
XX WPI; 2002-744015/81.
XX N-PSDB; ABX17290.
XX
XX Modification of a biodegradable polyester synthase, a mutant poly3-
PT hydroxybutanoate synthase, its preparation, a recombinant vector, a
PT transformant, preparation of a biodegradable ester polymer.
XX
XX Claim 13; Page 46-47; 124pp; Japanese.
XX
XX This invention relates to a novel method for the modification of an
CC enzyme participating to the biosynthesis of a poly3-hydroxyalkanoic acid
CC by modifying by recombinant DNA technology. The invention also comprises
CC a gene encoding the above mutant poly3-hydroxybutanoate synthase and a
CC recombinant vector containing the above gene. The method of the invention
CC may be used for the preparation of biodegradable polyesters. The present
CC sequence represents a protein used the method of the invention
XX
XX Sequence 589 AA;
SQ

Query Match 24.5%; Score 34; DB 5; Length 589;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 436 WYCWYLRHTYLQNE 449

RESULT 41
ABU10974
ID ABU10974 standard; protein; 589 AA.
XX
XX AC ABU10974;
XX
XX 04-FEB-2003 (first entry)
XX Poly3-hydroxybutanoate synthase mutant #1.
XX
XX Poly3-hydroxyalkanoic acid; biodegradable polyester; mutant; mutein.
XX
XX Synthetic.
XX
XX JP2002199890-A.
XX
XX 16-JUL-2002.
XX
XX 28-FEB-2001; 2001JP-00054717.
XX
XX 23-OCT-2000; 2000JP-00322748.
XX
XX (RIKA) RIKAGAKU KENKYUSHO.
XX

CC by modifying by recombinant DNA technology. The invention also comprises
 CC a gene encoding the above mutant poly3-hydroxybutanoate synthase and a
 CC recombinant vector containing the above gene. The method of the invention
 CC may be used for the preparation of biodegradable polyesters. The present
 CC sequence represents a protein used the method of the invention
 XX
 SQ Sequence 589 AA;
 Query Match 24.5%; Score 34; DB 5; Length 589;
 Best Local Similarity 42.9%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 Qy 39 WXXWXXXQTXLXNE 52
 Db 436 WYCWLRLRHTYLQNE 449
 RESULT 45
 ABU10989
 ID ABU10989 standard; protein; 589 AA.
 XX AC ABU10989;
 XX DT 04-FEB-2003 (first entry)
 XX DE Poly3-hydroxybutanoate synthase mutant associated protein #16.
 XX KW Poly3-hydroxyalkanoic acid; biodegradable polyester.
 XX OS Synthetic.
 XX PN JP2002199890-A.
 XX PD 16-JUL-2002.
 XX PF 28-FEB-2001; 2001JP-00054717.
 XX PR 23-OCT-2000; 2000JP-00322748.
 XX PA (RIKA) RIKAGAKU KENKYUSHO.
 XX DR WPI; 2002-744015/81.
 XX DR N-PSDB; ABX17305.
 XX PT Modification of a biodegradable polyester synthase, a mutant poly3-
 XX hydroxybutanoate synthase, its preparation, a recombinant vector, a
 XX transformant, preparation of a biodegradable ester polymer.
 XX PS Claim 13; Page 107-109; 124pp; Japanese.
 XX CC This invention relates to a novel method for the modification of an
 XX enzyme participating to the biosynthesis of a poly3-hydroxyalkanoic acid
 XX by modifying by recombinant DNA technology. The invention also comprises
 XX a gene encoding the above mutant poly3-hydroxybutanoate synthase and a
 XX recombinant vector containing the above gene. The method of the invention
 XX may be used for the preparation of biodegradable polyesters. The present
 XX sequence represents a protein used the method of the invention
 XX
 SQ Sequence 589 AA;
 Query Match 24.5%; Score 34; DB 5; Length 589;
 Best Local Similarity 42.9%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 Qy 39 WXXWXXXQTXLXNE 52
 Db 436 WYCWLRLRHTYLQNE 449
 RESULT 46
 ABU10962
 ID ABU10962 standard; protein; 589 AA.
 XX AC ABU10962;
 XX DT 04-FEB-2003 (first entry)
 XX DE Poly3-hydroxybutanoate synthase mutant associated protein #1.
 XX KW Poly3-hydroxyalkanoic acid; biodegradable polyester.
 XX OS Ralstonia eutropha.
 XX PN JP2002199890-A.
 XX PD 16-JUL-2002.
 XX PF 28-FEB-2001; 2001JP-00054717.
 XX PR 23-OCT-2000; 2000JP-00322748.
 XX PA (RIKA) RIKAGAKU KENKYUSHO.
 XX DR WPI; 2002-744015/81.
 XX DR N-PSDB; ABX17278.
 XX PT Modification of a biodegradable polyester synthase, a mutant poly3-
 XX hydroxybutanoate synthase, its preparation, a recombinant vector, a
 XX transformant, preparation of a biodegradable ester polymer.
 XX PS Claim 3; Page 15-16; 124pp; Japanese.
 XX CC This invention relates to a novel method for the modification of an
 XX enzyme participating to the biosynthesis of a poly3-hydroxyalkanoic acid
 XX by modifying by recombinant DNA technology. The invention also comprises
 XX a gene encoding the above mutant poly3-hydroxybutanoate synthase and a
 XX recombinant vector containing the above gene. The method of the invention
 XX may be used for the preparation of biodegradable polyesters. The present
 XX sequence represents a protein used the method of the invention
 XX
 SQ Sequence 589 AA;
 Query Match 24.5%; Score 34; DB 5; Length 589;
 Best Local Similarity 42.9%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 Qy 39 WXXWXXXQTXLXNE 52
 Db 436 WYCWLRLRHTYLQNE 449
 RESULT 47
 ABU10987
 ID ABU10987 standard; protein; 589 AA.
 XX AC ABU10987;
 XX DT 04-FEB-2003 (first entry)
 XX DE Poly3-hydroxybutanoate synthase mutant associated protein #14.
 XX KW Poly3-hydroxyalkanoic acid; biodegradable polyester.
 XX OS Synthetic.
 XX PN JP2002199890-A.
 XX PD 16-JUL-2002.
 XX PF 28-FEB-2001; 2001JP-00054717.
 XX PR 23-OCT-2000; 2000JP-00322748.
 XX PA (RIKA) RIKAGAKU KENKYUSHO.
 XX DR WPI; 2002-744015/81.

AC ABU10962;
 XX DT 04-FEB-2003 (first entry)
 XX DE Poly3-hydroxybutanoate synthase mutant associated protein #1.
 XX KW Poly3-hydroxyalkanoic acid; biodegradable polyester.
 XX OS Ralstonia eutropha.
 XX PN JP2002199890-A.
 XX PD 16-JUL-2002.
 XX PF 28-FEB-2001; 2001JP-00054717.
 XX PR 23-OCT-2000; 2000JP-00322748.
 XX PA (RIKA) RIKAGAKU KENKYUSHO.
 XX DR WPI; 2002-744015/81.
 XX DR N-PSDB; ABX17278.
 XX PT Modification of a biodegradable polyester synthase, a mutant poly3-
 XX hydroxybutanoate synthase, its preparation, a recombinant vector, a
 XX transformant, preparation of a biodegradable ester polymer.
 XX PS Claim 3; Page 15-16; 124pp; Japanese.
 XX CC This invention relates to a novel method for the modification of an
 XX enzyme participating to the biosynthesis of a poly3-hydroxyalkanoic acid
 XX by modifying by recombinant DNA technology. The invention also comprises
 XX a gene encoding the above mutant poly3-hydroxybutanoate synthase and a
 XX recombinant vector containing the above gene. The method of the invention
 XX may be used for the preparation of biodegradable polyesters. The present
 XX sequence represents a protein used the method of the invention
 XX
 SQ Sequence 589 AA;
 Query Match 24.5%; Score 34; DB 5; Length 589;
 Best Local Similarity 42.9%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 Qy 39 WXXWXXXQTXLXNE 52
 Db 436 WYCWLRLRHTYLQNE 449
 RESULT 47
 ABU10987
 ID ABU10987 standard; protein; 589 AA.
 XX AC ABU10987;
 XX DT 04-FEB-2003 (first entry)
 XX DE Poly3-hydroxybutanoate synthase mutant associated protein #14.
 XX KW Poly3-hydroxyalkanoic acid; biodegradable polyester.
 XX OS Synthetic.
 XX PN JP2002199890-A.
 XX PD 16-JUL-2002.
 XX PF 28-FEB-2001; 2001JP-00054717.
 XX PR 23-OCT-2000; 2000JP-00322748.
 XX PA (RIKA) RIKAGAKU KENKYUSHO.
 XX DR WPI; 2002-744015/81.

DR N-PSDB; ABX17303.

XX Modification of a biodegradable polyester synthase, a mutant poly3-hydroxybutanoate synthase, its preparation, a recombinant vector, a transformant, preparation of a biodegradable ester polymer.

PT Claim 13; Page 99-101; 124pp; Japanese.

XX This invention relates to a novel method for the modification of an enzyme participating to the biosynthesis of a poly3-hydroxyalkanoic acid by modifying by recombinant DNA technology. The invention also comprises a gene encoding the above mutant poly3-hydroxybutanoate synthase and a recombinant vector containing the above gene. The method of the invention may be used for the preparation of biodegradable polyesters. The present CC sequence represents a protein used the method of the invention

XX Sequence 589 AA;

Query Match 24.5%; Score 34; DB 5; Length 589;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNE 52
Db 436 WYCWLRLHTYLQNE 449

RESULT 48

ABU10981

ID ABU10981 standard; protein; 589 AA.

AC ABU10981;

XX 04-FEB-2003 (first entry)

XX Poly3-hydroxybutanoate synthase mutant #8.

XX Poly3-hydroxyalkanoic acid; biodegradable polyester; mutant; mutein.

XX Synthetic.

XX JP2002199890-A.

XX 16-JUL-2002.

XX 28-FEB-2001; 2001JP-00054717.

XX 23-OCT-2000; 2000JP-00322748.

XX (RIKA) RIKAGAKU KENKYUSHO.

XX WPI; 2002-744015/81.

XX N-PSDB; ABX17297.

XX Modification of a biodegradable polyester synthase, a mutant poly3-hydroxybutanoate synthase, its preparation, a recombinant vector, a transformant, preparation of a biodegradable ester polymer.

XX Claim 13; Page 75-76; 124pp; Japanese.

XX This invention relates to a novel method for the modification of an enzyme participating to the biosynthesis of a poly3-hydroxyalkanoic acid by modifying by recombinant DNA technology. The invention also comprises a gene encoding the above mutant poly3-hydroxybutanoate synthase and a recombinant vector containing the above gene. The method of the invention may be used for the preparation of biodegradable polyesters. The present CC sequence represents a protein used the method of the invention

XX Sequence 589 AA;

Query Match 24.5%; Score 34; DB 5; Length 589;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNE 52
Db 436 WYCWLRLHTYLQNE 449

Qy 39 WXXWXXXQTXLXNE 52
Db 436 WYCWLRLHTYLQNE 449

RESULT 49

ABU10986

ID ABU10986 standard; protein; 589 AA.

XX AC ABU10986;

XX 04-FEB-2003 (first entry)

XX Poly3-hydroxybutanoate synthase mutant associated protein #13.

XX Poly3-hydroxyalkanoic acid; biodegradable polyester.

XX Synthetic.

XX JP2002199890-A.

XX 16-JUL-2002.

XX 28-FEB-2001; 2001JP-00054717.

XX 23-OCT-2000; 2000JP-00322748.

XX (RIKA) RIKAGAKU KENKYUSHO.

XX WPI; 2002-744015/81.

XX N-PSDB; ABX17302.

XX Modification of a biodegradable polyester synthase, a mutant poly3-hydroxybutanoate synthase, its preparation, a recombinant vector, a transformant, preparation of a biodegradable ester polymer.

XX Claim 13; Page 95-97; 124pp; Japanese.

XX This invention relates to a novel method for the modification of an enzyme participating to the biosynthesis of a poly3-hydroxyalkanoic acid by modifying by recombinant DNA technology. The invention also comprises a gene encoding the above mutant poly3-hydroxybutanoate synthase and a recombinant vector containing the above gene. The method of the invention may be used for the preparation of biodegradable polyesters. The present CC sequence represents a protein used the method of the invention

XX Sequence 589 AA;

Query Match 24.5%; Score 34; DB 5; Length 589;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNE 52
Db 436 WYCWLRLHTYLQNE 449

RESULT 50

ABU10988

ID ABU10988 standard; protein; 589 AA.

XX AC ABU10988;

XX 04-FEB-2003 (first entry)

XX Poly3-hydroxybutanoate synthase mutant associated protein #15.

XX Poly3-hydroxyalkanoic acid; biodegradable polyester.

XX Synthetic.

XX JP2002199890-A.

XX 16-JUL-2002.
PD
XX
XX 28-FEB-2001; 2001JP-00054717.
PF
XX 23-OCT-2000; 2000JP-00322748.
PR
XX
XX (RIKA) RIKAGAKU KENKYUSHO.
PA
XX
XX WPI; 2002-744015/81.
DR
DR N-PSDB; ABX17304.
XX
XX Modification of a biodegradable polyester synthase, a mutant poly3-
PT hydroxybutanoate synthase, its preparation, a recombinant vector, a
PT transformant, preparation of a biodegradable ester polymer.
XX
XX Claim 13; Page 103-105; 124pp; Japanese.
PS
XX This invention relates to a novel method for the modification of an
CC enzyme participating to the biosynthesis of a poly3-hydroxyalkanoic acid
CC by modifying by recombinant DNA technology. The invention also comprises
CC a gene encoding the above mutant poly3-hydroxybutanoate synthase and a
CC recombinant vector containing the above gene. The method of the invention
CC may be used for the preparation of biodegradable polyesters. The present
CC sequence represents a protein used the method of the invention
XX
XX Sequence 589 AA;
SQ
Query Match 24.5%; Score 34; DB 5; Length 589;
Best Local Similarity 42.9%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 39 WXXWXXQTXLXNE 52
DB 436 WYCWYLRHTYLQNE 449
Search completed: September 30, 2004, 10:31:31
Job time : 58 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:32:37 ; Search time 133 Seconds
(without alignments)
157.270 Million cell updates/sec

Title: US-09-955-502-1

Perfect score: 139

Sequence: 1 MREXXXCXXXXXXXXXXXXX.....QTXLXNEXLXXXXXXXXXX 65

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*

3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*

7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*

10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*

11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*

12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*

13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*

14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*

15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*

16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*

17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*

18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	35.3	90	9	US-09-955-502-20
2	48	34.5	78	9	US-09-955-502-19
3	48	34.5	88	9	US-09-955-502-15
4	48	34.5	88	9	US-09-955-502-33
5	48	34.5	91	9	US-09-955-502-11
6	48	34.5	91	9	US-09-955-502-12
7	48	34.5	91	9	US-09-955-502-13
8	48	34.5	91	9	US-09-955-502-14
9	48	34.5	91	9	US-09-955-502-16
10	48	34.5	91	9	US-09-955-502-17
11	48	34.5	91	9	US-09-955-502-18
12	47	33.8	87	9	US-09-955-502-6
13	47	33.8	87	9	US-09-955-502-7
14	47	33.8	87	9	US-09-955-502-8
15	47	33.8	90	9	US-09-955-502-10

16	46	33.1	76	9	US-09-955-502-21	Sequence 21, Appl
17	46	33.1	88	9	US-09-955-502-9	Sequence 9, Appl
18	46	33.1	88	9	US-09-955-502-26	Sequence 26, Appl
19	46	33.1	88	9	US-09-955-502-27	Sequence 27, Appl
20	46	33.1	88	9	US-09-955-502-28	Sequence 28, Appl
21	46	33.1	90	9	US-09-955-502-23	Sequence 5, Appl
22	46	33.1	91	9	US-09-955-502-5	Sequence 32, Appl
23	45	32.4	87	9	US-09-955-502-24	Sequence 24, Appl
24	45	32.4	87	9	US-09-955-502-29	Sequence 29, Appl
25	45	32.4	87	9	US-09-955-502-30	Sequence 30, Appl
26	45	32.4	87	9	US-09-955-502-31	Sequence 31, Appl
27	45	32.4	87	9	US-09-955-502-32	Sequence 32, Appl
28	44	31.7	86	9	US-09-955-502-4	Sequence 4, Appl
29	44	31.7	87	9	US-09-955-502-2	Sequence 2, Appl
30	44	31.7	87	9	US-09-955-502-3	Sequence 3, Appl
31	44	31.7	87	9	US-09-955-502-25	Sequence 25, Appl
32	44	31.7	89	9	US-09-955-502-22	Sequence 22, Appl
33	42	30.2	65	9	US-09-955-502-1	Sequence 1, Appl
34	34	24.5	356	16	US-10-357-521-5	Sequence 5, Appl
35	34	24.5	549	15	US-10-359-493-13068	Sequence 13068, A
36	33	23.7	391	15	US-10-359-493-8282	Sequence 8282, AP
37	33	23.7	393	14	US-10-156-761-14669	Sequence 14669, A
38	33	23.7	422	12	US-10-282-122A-61450	Sequence 61450, A
39	33	23.7	1235	12	US-10-282-122A-60959	Sequence 60959, A
40	32	23.0	388	12	US-10-425-114-59697	Sequence 59697, A
41	32	23.0	539	12	US-10-424-599-276981	Sequence 276981, A
42	32	23.0	1181	16	US-10-406-765A-21177	Sequence 2177, AP
43	31	22.3	91	12	US-10-424-599-226079	Sequence 226079, A
44	31	22.3	311	16	US-10-437-963-184047	Sequence 184047, A
45	31	22.3	372	16	US-10-437-963-145960	Sequence 145960, A
46	31	22.3	384	16	US-10-437-963-119949	Sequence 119949, A
47	31	22.3	398	12	US-10-425-114-53395	Sequence 53395, A
48	31	22.3	398	12	US-10-425-114-64162	Sequence 64162, A
49	31	22.3	398	12	US-10-425-114-72211	Sequence 72211, A
50	31	22.3	584	16	US-10-437-963-109347	Sequence 109347, A

ALIGNMENTS

RESULT 1

US-09-955-502-20

; Sequence 20, Application US/09955502

; Patent No. US20020072118A1

; GENERAL INFORMATION:

; APPLICANT: Downs, Diana M.

; APPLICANT: Gralnick, Jeff A.

; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and

; FILE REFERENCE: 960296.97559

; CURRENT APPLICATION NUMBER: US/09/955,502

; CURRENT FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: 60/234,588

; PRIOR FILING DATE: 2000-09-22

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 20

; TYPE: PRT

; ORGANISM: Versinia pestis

US-09-955-502-20

Query Match 35.3%; Score 49; DB 9; Length 90;

Best local Similarity 23.6%; Pred. No. 0.0071;

Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MREXXXCXXXXXXXXXXXXX.....QTXLXNEXL 55

Db 1 MSRTIFCTFLKDAERQDFQLYGEIGKRYNEISKEANSQWITKQTMLINEKL 55

RESULT 2

US-09-955-502-19

; Sequence 19, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-955-502-19

Query Match 34.5%; Score 48; DB 9; Length 78;
Best Local Similarity 23.6%; Pred. No. 0.01;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXCCXXXXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTLXNEXXL 55
Db 1 MSRTIFCTFLQREAGDQDFQYFGLGKRIYNEISKEAWAQWQHKTMLINEKKL 55

RESULT 3

US-09-955-502-15
; Sequence 15, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Salmonella enteritidis
US-09-955-502-15

Query Match 34.5%; Score 48; DB 9; Length 88;
Best Local Similarity 23.6%; Pred. No. 0.01;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXCCXXXXXXXXXXXXXPPXXXXXXXXXXXXXQTLXNEXXL 55
Db 1 MSRTIFCTFLQREAGDQDFQYFGLGKRIYNEISKEAWAQWQHKTMLINEKKL 55

RESULT 4

US-09-955-502-33
; Sequence 33, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588

; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Coxiella burnetii
US-09-955-502-33

Query Match 34.5%; Score 48; DB 9; Length 88;
Best Local Similarity 23.6%; Pred. No. 0.01;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXCCXXXXXXXXXXXXXPPXXXXXXXXXXXXXQTLXNEXXL 55
Db 1 MTRRIICQKLGKADALNYSYPGELGERIYNHISEQAWQAWLSHOTMLINEYEL 55

RESULT 5

US-09-955-502-11
; Sequence 11, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Escherichia coli K-12 MG1655
US-09-955-502-11

Query Match 34.5%; Score 48; DB 9; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.01;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXCCXXXXXXXXXXXXXPPXXXXXXXXXXXXXQTLXNEXXL 55
Db 1 MSRTIFCTFLQREAGDQDFQYFGLGKRIYNEISKEAWAQWQHKTMLINEKKL 55

RESULT 6

US-09-955-502-12
; Sequence 12, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Escherichia coli O157:H7EDL933
US-09-955-502-12

Query Match 34.5%; Score 48; DB 9; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.01;

```
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 MRRXXXCXXXXXXXXXXXXPXXXXXXXXXXXXXXXXXXXXXXXXXQTXLXNEXKL 55
Db 1 MSRTIFCFLQREAGQDFQLYPGELGKRIYNEISKDAWAQWQHQTMLINEKXL 55

RESULT 7
US-09-955-502-13
; Sequence 13, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Salmonella dublin
US-09-955-502-16
Query Match 34.5%; Score 48; DB 9; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 MRRXXXCXXXXXXXXXXXXPXXXXXXXXXXXXXXXXXXXXXXXXXQTXLXNEXKL 55
Db 1 MSRTIFCFLQREAGQDFQLYPGELGKRIYNEISKDAWAQWQHQTMLINEKXL 55

RESULT 8
US-09-955-502-14
; Sequence 14, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Salmonella paratyphi
US-09-955-502-14
Query Match 34.5%; Score 48; DB 9; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 MRRXXXCXXXXXXXXXXXXPXXXXXXXXXXXXXXXXXXXXXXXXXQTXLXNEXKL 55
Db 1 MSRTIFCFLQREAGQDFQLYPGELGKRIYNEISKDAWAQWQHQTMLINEKXL 55

RESULT 9
US-09-955-502-16
; Sequence 16, Application US/09955502
; Patent No. US20020072118A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Salmonella dublin
US-09-955-502-16
Query Match 34.5%; Score 48; DB 9; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 MRRXXXCXXXXXXXXXXXXPXXXXXXXXXXXXXXXXXXXXXXXXXQTXLXNEXKL 55
Db 1 MSRTIFCFLQREAGQDFQLYPGELGKRIYNEISKDAWAQWQHQTMLINEKXL 55

RESULT 10
US-09-955-502-17
; Sequence 17, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Salmonella typhi CT18
US-09-955-502-17
Query Match 34.5%; Score 48; DB 9; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.012;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 MRRXXXCXXXXXXXXXXXXPXXXXXXXXXXXXXXXXXXXXXXXXXQTXLXNEXKL 55
Db 1 MSRTIFCFLQREAGQDFQLYPGELGKRIYNEISKDAWAQWQHQTMLINEKXL 55

RESULT 11
US-09-955-502-18
; Sequence 18, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; NUMBER OF SEQ ID NOS: 33
```



```
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Buchnera sp. APS
US-09-955-502-21

Query Match      33.1%; Score 46; DB 9; Length 76;
Best Local Similarity 23.6%; Pred. No. 0.027;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy      1 MRRXXCCXXXXXXXXXXXXXXXPPXXXXXXXQXXXXXXXXXXQTXLNEXXL 55
Db      1 MRRIFCTFFKKKGGQDFQSPGKLGKIKYDQISKXAKWKWKQILINEENL 55

RESULT 17
US-09-955-502-9
; Sequence 9, Application US/099555502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Shewanella putrefaciens
US-09-955-502-9

Query Match      33.1%; Score 46; DB 9; Length 88;
Best Local Similarity 23.6%; Pred. No. 0.031;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy      1 MRRXXCCXXXXXXXXXXXXXXXPPXXXXXXXQXXXXXXXXXXQTXLNEXXL 55
Db      1 MARTVNCVHLNKEADGLDFQLYPGDLGKRIFDNIKSEAWGLWQKKQTMLINEKXL 55

RESULT 18
US-09-955-502-26
; Sequence 26, Application US/099555502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
```

```
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-09-955-502-26

Query Match      33.1%; Score 46; DB 9; Length 88;
Best Local Similarity 30.3%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy      23 PXXGXXXXXXXQXXXXXXXXXXQTXLNEXXL 55
Db      23 PNELGKRIFENVSQEAWAAWTRHQTMLINENL 55

RESULT 19
US-09-955-502-27
; Sequence 27, Application US/099555502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Neisseria meningitidis B
US-09-955-502-27

Query Match      33.1%; Score 46; DB 9; Length 88;
Best Local Similarity 30.3%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy      23 PXXGXXXXXXXQXXXXXXXXXXQTXLNEXXL 55
Db      23 PNELGKRIFENVSQEAWAAWTRHQTMLINENL 55

RESULT 20
US-09-955-502-28
; Sequence 28, Application US/099555502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Neisseria meningitidis A
US-09-955-502-28

Query Match      33.1%; Score 46; DB 9; Length 88;
Best Local Similarity 30.3%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy      23 PXXGXXXXXXXQXXXXXXXXXXQTXLNEXXL 55
```

Db 23 PNELGKRIPEVNSQEAWAATRHQTMLINERL 55

RESULT 21

US-09-955-502-23
; Sequence 23, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-09-955-502-23

Query Match 33.1%; Score 46; DB 9; Length 90;
Best Local Similarity 23.6%; Pred. No. 0.031;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MVRXXKXXXXXXXPPXXGXXXXXXXQTXLXNEXXL 55
Db 1 MTRVCMCRYKEELPLGLERAPYFCAGKEDIFNHVSQKAWADWQKHQTMLINERL 55

RESULT 22

US-09-955-502-5
; Sequence 5, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Actinobacillus actinomycetemcomitans
US-09-955-502-5

Query Match 33.1%; Score 46; DB 9; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.031;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MVRXXKXXXXXXXPPXXGXXXXXXXQTXLXNEXXL 55
Db 1 MARMVFCERLQKQEAEGDLFQLYGEIGKRFIDISIKQANGENWKKQTMLVNEKXL 55

RESULT 23

US-09-955-502-24
; Sequence 24, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and

; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-09-955-502-24

Query Match 32.4%; Score 45; DB 9; Length 87;
Best Local Similarity 30.3%; Pred. No. 0.05;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 23 PXXGXXXXXXXQTXLXNEXXL 55
Db 23 PGAKGQDIFEHISQKAWADWQKHQTMLINERL 55

RESULT 24

US-09-955-502-29
; Sequence 29, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Burkholderia mallei
US-09-955-502-29

Query Match 32.4%; Score 45; DB 9; Length 87;
Best Local Similarity 23.6%; Pred. No. 0.05;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MVRXXKXXXXXXXPPXXGXXXXXXXQTXLXNEXXL 55
Db 1 MARMHCAKLGKQEAEGDLFPPLFGLGKRLYESVSKQAWQDLKQQTMLINERL 55

RESULT 25

US-09-955-502-30
; Sequence 30, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 87
; TYPE: PRT

```
; ORGANISM: Burkholderia pseudomallei
US-09-955-502-30

Query Match      32.4%; Score 45; DB 9; Length 87;
Best Local Similarity 23.6%; Pred. No. 0.05;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXKXXXXXXXXXXXXXPPXGXXXXXQXXXXXXXXXXQTXLNEXXL 55
Db 1 MARMTHCAKLGKEAGSLDFPPLPGELGKELYESVSKQAWQDLKQOTMLINL 55

RESULT 26
US-09-955-502-31
; Sequence 31, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Thibacillus ferrooxidans
US-09-955-502-31

Query Match      32.4%; Score 45; DB 9; Length 87;
Best Local Similarity 30.3%; Pred. No. 0.05;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 23 PXXGXXXXXXXQXXQXXQXXQXXQXXQXXQXXQXXQXXQXXQXX 55
Db 23 PGALGARIYQEVSKAWQGLKHQOTMLINEYRL 55

RESULT 27
US-09-955-502-32
; Sequence 32, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Methylococcus capsulatus
US-09-955-502-32

Query Match      32.4%; Score 45; DB 9; Length 87;
Best Local Similarity 22.2%; Pred. No. 0.05;
Matches 14; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 1 MRRXXKXXXXXXXXXXXXXPPXGXXXXXQXXXXXXXXXXQTXLNEXXL 60
Db 1 MARRIICAKLGTEAGDLAPFPFGQGORIFEHVSKEAWQDLKLOTMLINEHRLTFPEA 60
```

```
Qy 61 XXR 63
Db 61 SAR 63

RESULT 28
US-09-955-502-4
; Sequence 4, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Bordetella bronchiseptica
US-09-955-502-4

Query Match      31.7%; Score 44; DB 9; Length 86;
Best Local Similarity 26.8%; Pred. No. 0.081;
Matches 11; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 23 PXXGXXXXXXXQXXQXXQXXQXXQXXQXXQXXQXXQXXQXXQXX 63
Db 23 PGELGTRIWOQISKEAWEWKQIQITLVNENLNLADARAR 63

RESULT 29
US-09-955-502-2
; Sequence 2, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-09-955-502-2

Query Match      31.7%; Score 44; DB 9; Length 87;
Best Local Similarity 26.8%; Pred. No. 0.082;
Matches 11; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 23 PXXGXXXXXXXQXXQXXQXXQXXQXXQXXQXXQXXQXXQXXQXX 63
Db 23 PGELGTRIWOQISKEAWEWKQIQITLVNENLNLADARAR 63

RESULT 30
US-09-955-502-3
; Sequence 3, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
```


FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (53)..(54)
; OTHER INFORMATION: can be any amino acid
FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (56)..(62)
; OTHER INFORMATION: can be any amino acid
FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (64)..(65)
; OTHER INFORMATION: can be any amino acid
US-09-955-502-1

Query Match 30.2%; Score 42; DB 9; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.18; Indels 0;
Matches 17; Conservative 0; Mismatches 0; Gaps 0;

QY 39 WXXWXXQTXLNEXXL 55
DB 39 WXXWXXQTXLNEXXL 55

RESULT 34

US-10-357-521-5
; Sequence 5, Application US/10357521
; Publication No. US20040088751A1
GENERAL INFORMATION:

APPLICANT: LIEBERGSELL, MATTHIAS
STEINBUCHER, ALEXANDER
TITLE OF INVENTION: PRODUCTION OF POLYALKANOATE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: DC

COUNTRY: USA
ZIP: 20005-3518
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/357,521
FILING DATE: 03-Feb-2003
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/178,257
FILING DATE: 11-JAN-1994
APPLICATION NUMBER: PCT/GB92/01291
FILING DATE: 15-JUL-1992
APPLICATION NUMBER: GB 9115245.4
FILING DATE: 16-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 9320/206058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/861-3000
TELEFAX: 202/822-0944
TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Alcaligenes eutrophus
SEQUENCE DESCRIPTION: SEQ ID NO: 5;

US-10-357-521-5

Query Match 24.5%; Score 34; DB 16; Length 356;
Best Local Similarity 42.9%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXQTXLXNE 52
DB 264 WYCWYLRHTYLQNE 277

RESULT 35

US-10-369-493-13068
; Sequence 13068, Application US/10369493
; Publication No. US20030233675A1
GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 13068
LENGTH: 549
TYPE: PRT

ORGANISM: Aspergillus nidulans
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(549)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-13068

Query Match 24.5%; Score 34; DB 15; Length 549;
Best Local Similarity 42.9%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXQTXLXNE 52
DB 88 WYWLWTVLVANE 101

RESULT 36

US-10-369-493-8282
; Sequence 8282, Application US/10369493
; Publication No. US20030233675A1
GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 8282
LENGTH: 391
TYPE: PRT

ORGANISM: Thermobifida fusca
US-10-369-493-8282

Query Match 23.7%; Score 33; DB 15; Length 391;
Best Local Similarity 54.5%; Pred. No. 61;

Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXTXLXNE 52
| | | | |
DB 218 WKVQTTLMNE 228

RESULT 37

US-10-156-761-14669
; Sequence 14669, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14669
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14669

Query Match 23.7%; Score 33; DB 14; Length 393;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXTXLXNE 52
| | | | |
DB 220 WRVQTTLMNE 230

RESULT 38

US-10-282-122A-61450
; Sequence 61450, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1

Query Match 23.7%; Score 33; DB 14; Length 393;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61450
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Legionella pneumophila
US-10-282-122A-61450

Query Match 23.7%; Score 33; DB 12; Length 422;
Best Local Similarity 35.3%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 39 WXXXTXLXNEXXL 55
| | | | |
DB 292 WSGWKQLTTLGNKVL 308

RESULT 39

US-10-282-122A-60959
; Sequence 60959, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 60959
; LENGTH: 1235
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60959

Query Match      23.7%; Score 33; DB 12; Length 1235;
Best Local Similarity 30.8%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXN 51
Db 276 WSSWTHLKTSTEN 288

RESULT 40
US-10-425-114-59697
; Sequence 59697, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59697
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-231-B7_FLI.pep
US-10-425-114-59697

Query Match      23.0%; Score 32; DB 12; Length 388;
Best Local Similarity 35.7%; Pred. No. 99;
Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNE 52
Db 224 WALMTVLQSAIIKE 237

RESULT 41
US-10-424-599-276981
; Sequence 276981, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 276981
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_92132C.1.pep
US-10-424-599-276981

Query Match      23.0%; Score 32; DB 12; Length 539;
```

```
; Best Local Similarity 23.5%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNEXXL 55
Db 210 WSAWSARESAIEQEKLL 226

RESULT 42
US-10-408-765A-2177
; Sequence 2177, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Koin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Wartock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2177
; LENGTH: 1191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2177

Query Match      23.0%; Score 32; DB 16; Length 1191;
Best Local Similarity 38.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXN 51
Db 337 WAPWGSSTHLDN 349

RESULT 43
US-10-424-599-226079
; Sequence 226079, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 226079
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_4617C.1.pep
US-10-424-599-226079

Query Match      22.3%; Score 31; DB 12; Length 91;
Best Local Similarity 35.3%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNEXXL 55
Db 18 WEHWYSTQTRWQENRL 34
```

```
RESULT 44
US-10-437-963-184047
; Sequence 184047, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 184047
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81078C.1.pep
US-10-437-963-184047

Query Match      22.3%; Score 31; DB 16; Length 311;
Best Local Similarity 35.7%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      39 WXXWXXXQTXLXN 51
DB      246 WAKWLDEQKLAN 258

RESULT 45
US-10-437-963-145960
; Sequence 145960, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 145960
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_46630C.1.pep
US-10-437-963-145960

Query Match      22.3%; Score 31; DB 16; Length 372;
Best Local Similarity 35.7%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      39 WXXWXXXQTXLXNE 52
DB      194 WSLWLVLSPLNNE 207

RESULT 46
US-10-437-963-119949
; Sequence 119949, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 119949
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_23115C.1.pep
US-10-437-963-119949

Query Match      22.3%; Score 31; DB 16; Length 384;
Best Local Similarity 35.7%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      39 WXXWXXXQTXLXNE 52
DB      196 WSLWLVLSQSLKE 209

RESULT 47
US-10-425-114-53395
; Sequence 53395, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53395
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700451509_FLI.pep
US-10-425-114-53395

Query Match      22.3%; Score 31; DB 12; Length 398;
Best Local Similarity 45.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      39 WXXWXXXQTXL 49
DB      303 WGSLSLFQILL 313

RESULT 48
US-10-425-114-64162
; Sequence 64162, Application US/10425114
```

Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 64162
LENGTH: 398
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3592-097-D6_FLI.pep
US-10-425-114-64162
Query Match 22.3%; Score 31; DB 12; Length 398;
Best Local Similarity 45.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXL 49
DB 303 WGWWSLFQTL 313
RESULT 49
US-10-425-114-72211
Sequence 72211, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 72211
LENGTH: 398
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700428946_FLI.pep
US-10-425-114-72211
Query Match 22.3%; Score 31; DB 12; Length 398;
Best Local Similarity 45.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXL 49
DB 303 WGWWSLFQTL 313
RESULT 50
US-10-437-963-109347
Sequence 109347, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 109347
LENGTH: 984
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_13512C.1.pep
US-10-437-963-109347
Query Match 22.3%; Score 31; DB 16; Length 984;
Best Local Similarity 45.5%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXL 49
DB 375 WLEWILSQTL 385
Search completed: September 30, 2004, 10:44:09
Job time : 134 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:29:36 ; Search time 16 Seconds
(without alignments)
390.778 Million cell updates/sec

Title: US-09-955-502-1
Perfect score: 139
Sequence: 1 MXXXXXXXXXXXXXXXXXXXXX.....QTXLXNEXXLLXXXXXXXXXX 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : PIR_78.*

- 1: Pirl.*
- 2: Pirl2.*
- 3: Pirl3.*
- 4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49	35.3	90	2 A10116	conserved hypoteth
2	48	34.5	91	2 A65082	hypothetical prote
3	48	34.5	91	2 A5954	hypothetical prote
4	48	34.5	91	2 A10879	conserved hypoteth
5	48	34.5	91	2 F91108	hypothetical prote
6	47	33.8	90	2 C64013	hypothetical prote
7	47	33.8	90	2 C82320	conserved hypoteth
8	46	33.1	88	2 H81014	conserved hypoteth
9	46	33.1	93	2 E84994	hypothetical prote
10	44	31.7	90	2 H83003	conserved hypoteth
11	44	31.7	105	2 C82624	conserved hypoteth
12	34	24.5	589	2 A34341	poly(3-hydroxybuty
13	33	23.7	564	2 F90965	hypothetical prote
14	33	23.7	564	2 F85813	hypothetical prote
15	33	23.7	569	2 H64959	probable membrane
16	33	23.7	1235	2 AC1358	ATP-dependent deox
17	32	23.0	380	2 T8251	ORF MSV090 probabl
18	32	23.0	420	2 S77102	hypothetical prote
19	32	23.0	1194	2 C59436	KIAA1391 protein [
20	32	23.0	3655	2 T38084	TRAP-like protein
21	31	22.3	211	2 T10392	late expression fa
22	31	22.3	567	2 AC0754	probable membrane
23	31	22.3	593	2 C64097	probable soluble 1
24	30	21.6	99	2 F91126	hypothetical prote
25	30	21.6	99	2 E85971	hypothetical prote
26	30	21.6	99	2 A65099	hypothetical prote
27	30	21.6	99	2 AE0895	conserved hypoteth
28	30	21.6	331	1 C63026	acetylpolymaine am
29	30	21.6	331	2 T04525	hypothetical prote

ALIGNMENTS

RESULT 1

A10116 conserved hypothetical protein YP00953 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Dec-2002
C:Accession: A10116
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, I. Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: A10116
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC89796.1; PID:gl15979022; GSPDB:GNO0175
C:Genetics:
A:Gene: YP00953
C:Superfamily: oxygen-labile Fe-S cluster protectant

Query Match 35.3%; Score 49; DB 2; Length 90;
Best Local Similarity 23.6%; Pred. No. 0.0022;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MXXXXXXXXXXXXXXXXXXXXXFXXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db 1 MSRTIFCTFLKDAERQDFQLYPFGTGRVNEISKEAWSQMTTKQTMLINEXKL 55

RESULT 2

A65082 hypothetical protein b2962 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Dec-2002
C:Accession: A65082
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coli .A.; Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: A65082
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-91 <BLAT>
A:Cross-references: GB:AE000378; GB:U00096; PIDN:AAC75999.1; PID:gl1789332.
A:Experimental source: strain K-12, substrain MG1655

RESULT 7

C82320
conserved hypothetical protein VC0451 [imported]; - Vibrio cholerae (strain N16961 serogroup O1)

C.Species: Vibrio cholerae
C.Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Dec-2002
C.Accession: C82320

A.Heldelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; L.R.R.; Mekalanos, D.; Esmalaeva, M.D.; Vamathevan, J.; Bacs, S.; Qin, H.; Dragoi, I.; Sellers, P.; Venter, J.C.; Fraser, C.M.

Feature 406, 477-483, 2000

A.Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A.Reference number: A82035; UID:20406833; PMID:10952301
A.Accession: C82320
A.Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-90 <HBI>
 A:Cross-references: GB:AE004132; GB:AE003852; NID:G9654871; PIDN:AAF93624.1; GSPDB:GN001
 A:Experimental source: serogroup O1, strain N16981; biotype EI Tor
 C:Genetics:
 A:Gene: VC0451
 A:Map position: 1
 C:Superfamily: oxygen-labile Fe-S cluster protectant

Query Match 33.8%; Score 47; DB 2; Length 90;
 Best Local Similarity 23.6%; Pred. No. 0.0057;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 MXPXXXCXNXXXXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLNEXXL 55
 DB 1 MARTVFCTRLQKADGLDFQYPGELGKRFIDNICKKAWAQWQTKQTLNKKEL 55

RESULT 8
 H81014
 Conserved hypothetical protein NMB2021 [imported] - Neisseria meningitidis (strain MC58)
 C:Species: Neisseria meningitidis
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Dec-2002
 C:Accession: H81014; F81958
 R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
 Science 287, 189-195, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: H81014
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-88 <FT>
 A:Cross-references: GB:AE002552; GB:AE002098; NID:G7227279; PIDN:AAF42344.1; PID:G722728
 A:Experimental source: serogroup B, strain MC58
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: F81958
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-88 <PAR>
 A:Cross-references: GB:AL162753; GB:AL157959; NID:G7379120; PIDN:CAB83718.1; PID:G737916
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: NMB2021; NMA0419
 C:Superfamily: oxygen-labile Fe-S cluster protectant

Query Match 33.1%; Score 46; DB 2; Length 88;
 Best Local Similarity 30.3%; Pred. No. 0.009;
 Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 23 PXXGXXXXXXXNXXXXXXXXXXXXXQTXLNEXXL 55
 DB 23 PNLGKRFENVSQBAWAAWTRHQIWLINENL 55

RESULT 9
 E84994
 Hypothetical protein [imported] - Buchnera sp. (strain APS)
 C:Species: Buchnera sp.
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Dec-2002
 C:Accession: E84994
 R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000
 A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
 A:Reference number: A84930; MUID:20445173; PMID:10993077
 A:Accession: E84994
 A:Status: Preliminary

A:Molecule type: DNA
 A:Residues: 1-93 <STO>
 A:Cross-references: GB:AP00398; GSPDB:GN00144
 A:Experimental source: strain APS
 C:Genetics:
 A:Gene: yggX; BU553
 C:Superfamily: oxygen-labile Fe-S cluster protectant

Query Match 33.1%; Score 46; DB 2; Length 93;
 Best Local Similarity 23.6%; Pred. No. 0.0094;
 Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 MRRXXXXXNXXXXXXXXXXXXXPPXXXXXXXXXXXXXQTXLNEXXL 55
 DB 17 MNRIFCTFFKKKSEQDFQSYPGKLGKIKYDQISKAWKWKIQTILINENL 71

RESULT 10
 H83003
 Conserved hypothetical protein PAS148 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Dec-2002
 C:Accession: H83003
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bro
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10994043
 A:Accession: H83003
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-90 <STO>
 A:Cross-references: GB:AE004927; GB:AE004091; NID:G9951437; PIDN:AAG08533.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PAS148
 C:Superfamily: oxygen-labile Fe-S cluster protectant

Query Match 31.7%; Score 44; DB 2; Length 90;
 Best Local Similarity 30.3%; Pred. No. 0.024;
 Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 23 PXXGXXXXXXXNXXXXXXXXXXXXXQTXLNEXXL 55
 DB 23 PGAKGEDIYNNVSRKAWDEWQKQHTMLINERRL 55

RESULT 11
 C82624
 Conserved hypothetical protein XF1908 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Dec-2002
 C:Accession: C82624
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: C82624
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-105 <SIM>
 A:Cross-references: GB:AE004010; GB:AE003849; NID:G9106992; PIDN:AAF84714.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
 Submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.,

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A.; Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1908
C;Superfamily: oxygen-labile Fe-S cluster protectant

Query Match 31.7%; Score 44; DB 2; Length 105;
Best Local Similarity 47.1%; Pred. No. 0.028; 9; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 9;

QY 39 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 54 WAAMLVHQTMLNENRL 70
| | | | |

RESULT 12

A34341
poly(3-hydroxybutyrate) synthase (EC 2.3.1.-) - Alcaligenes eutrophus
C;Species: Alcaligenes eutrophus
C;Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 08-Sep-2000
C;Accession: A34341; A39190
R;Peoples, O.P.; Sinskey, A.J.
J. Biol. Chem. 264, 15298-15303, 1989
A;Title: Poly-beta-hydroxybutyrate (PHB) biosynthesis in Alcaligenes eutrophus H16. Ident
A;Reference number: A34341; MUID:89359357; PMID:2670936
A;Accession: A34341
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-589 <PEO>
A;Cross-references: GB:J05003; NID:gl41958; PIDN:AAA21975.1; PID:gl41959
A;Experimental source: strain H16
R;Schubert, P.; Krueger, N.; Steinbuechel, A.
J. Bacteriol. 173, 168-175, 1991
A;Title: Molecular analysis of the Alcaligenes eutrophus poly(3-hydroxybutyrate) biosyn-
omoter.
A;Reference number: A39190; MUID:91100279; PMID:1987116
A;Accession: A39190
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-219 <SCH>
A;Cross-references: GB:M64341; NID:gl41964; PIDN:AAA21979.1; PID:gl41965
A;Note: the authors translated the codon TAC for residue 120 as thr
C;Superfamily: poly(3-hydroxyalkanoic acid) synthase phbc
C;Keywords: acyltransferase

Query Match 24.5%; Score 34; DB 2; Length 589;
Best Local Similarity 42.9%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 436 WYCWYLRHTYQNE 449
| | | | |

RESULT 13

F90965
hypothetical protein EC2694 [imported] - Escherichia coli (strain O157:H7, substrain R1
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: F90965
R;Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90965
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-564 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA836117.1; PID:gl3362162; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain R1MD 0509952
C;Genetics:
A;Gene: EC2694

Query Match 23.7%; Score 33; DB 2; Length 564;
Best Local Similarity 35.3%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 396 WQAWHDTLTRYNREGAL 412
| | | | |

RESULT 14

F85813
hypothetical protein Z3047 [imported] - Escherichia coli (strain O157:H7, substrain EDL93;
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: F85813
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimaanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85813
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-564 <STO>
A;Cross-references: GB:AE005174; NID:gl2516062; PIDN:AAG56970.1; GSPDB:GN00145; UWGP:Z304;
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3047

Query Match 23.7%; Score 33; DB 2; Length 564;
Best Local Similarity 35.3%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 396 WQAWHDTLTRYNREGAL 412
| | | | |

RESULT 15

H64959
probable membrane protein b1956 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: H64959
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64959
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-569 <BLAT>
A;Cross-references: GB:AE000287; GB:U00096; NID:gl788257; PIDN:AAC75022.1; PID:gl788266;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Start codon: TTG
C;Keywords: transmembrane protein
F;30-46/Domain: transmembrane #status predicted <TM1>
F;363-379/Domain: transmembrane #status predicted <TM2>

Query Match 23.7%; Score 33; DB 2; Length 569;
Best Local Similarity 35.3%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 401 WQAWHDTLTRYNREGAL 417
| | | | |

RESULT 16

AC1358 ATP-dependent deoxyribonuclease (chain A) homolog lmo2267 [imported] - *Listeria monocytogenes*C/Species: *Listeria monocytogenes*

C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C/Accession: AC1358

R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A/Authors: Kreft, J.; Kuhn, M.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluster, T.; Simoes, N.; Rierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.

A/Title: Comparative genomics of *Listeria* species.

A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AC1358

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1235 <GLA>

A/Cross-references: GB:NC_003210; PIDN:CAD00345.1; PID:gl6411737; GSPDB:GN00177

A/Experimental source: strain EGD-e

C/Genetics:

A/Gene: lmo2267

Query Match 23.7%; Score 33; DB 2; Length 1235;

Best Local Similarity 30.8%; Pred. No. 53;

Matches 4; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 39 WXXWXXQTXLXN 51

Db 276 WSSWTHLKTSTEN 288

RESULT 17

T28251 ORF MSV090 Probable Molluscum contagiosum virus MCL21L (vaccinia A16L) homolog - *Melanoplus sanguinipes*C/Species: *Melanoplus sanguinipes* entomopoxvirus

C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C/Accession: T28251

R/Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.

J. Virol. 73, 533-552, 1999

A/Title: The genome of *Melanoplus sanguinipes* entomopoxvirus.

A/Reference number: Z20484; MUID:99102612; PMID:9847359

A/Accession: T28251

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-380 <AFO>

A/Cross-references: EMBL:AF063866; NID:g4049647; PIDN:AAC97640.1; PID:g4049680

C/Genetics:

A/Note: MSV090

Query Match 23.0%; Score 32; DB 2; Length 380;

Best Local Similarity 42.9%; Pred. No. 29;

Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 42 WXXWXXQTXLXNXXL 55

Db 186 WLESQTXLNDIAL 199

RESULT 18

S77102 hypothetical protein slr1865 - *Synechocystis* sp. (strain PCC 6803)C/Species: *Synechocystis* sp.

A/Variety: PCC 6803

C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C/Accession: S77102

R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

s.

Query Match 23.0%; Score 32; DB 2; Length 3655;

A/Reference number: S74322; MUID:97061201; PMID:8905231

A/Accession: S77102

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-420 <KAN>

A/Cross-references: EMBL:D90908; GB:AB001339; NID:gl652725; PIDN:BAAL7660.1; PID:d101839;

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C/Genetics:

A/Start codon: GTG

Query Match 23.0%; Score 32; DB 2; Length 420;

Best Local Similarity 35.7%; Pred. No. 32;

Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXQTXLXNE 52

Db 150 WRFWEVAQTNLKQ 163

RESULT 19

C59436 KIAA1391 protein [imported] - human

C/Species: *Homo sapiens* (man)

C/Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 23-Sep-2002

C/Accession: C59436; D59436

R/Nagase, T.; Kikuno, R.; Ishikawa, K.I.; Hirose, M.; Ohara, O.

DNA Res. 7, 65-73, 2000

A/Title: Prediction of the coding sequences of unidentified human genes. XVI. The complet

A/Reference number: C59436

A/Accession: C59436

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1194 <NAG>

A/Cross-references: GB:BAA92629; PID:g7243163; PIDN:BAA92629.1

R/Ohara, O.; Nagase, T.; Kikuno, R.

submitted to GenBank, January 2000

A/Reference number: D59436

A/Accession: D59436

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1194 <OHA>

A/Cross-references: GB:BAA92629; PID:g7243163; PIDN:BAA92629.1

Query Match 23.0%; Score 32; DB 2; Length 1194;

Best Local Similarity 38.5%; Pred. No. 83;

Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXQTXLXN 51

Db 340 WAFWRGSSSTHLDN 352

RESULT 20

T38084

TRAP-like protein - fission yeast (*Schizosaccharomyces pombe*)C/Species: *Schizosaccharomyces pombe*

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C/Accession: T38084

R/Gentles, S.; Odell, C.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, November 1995

A/Reference number: Z21768

A/Accession: T38084

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-3655 <GEN>

A/Cross-references: EMBL:Z68136; PIDN:CAA92239.1; GSPDB:GN000066; SPDB:SPAC1F5.11C

A/Experimental source: strain 972h-, cosmid c1F5

C/Genetics:

A/Gene: SPDB:SPAC1F5.11c

A/Map position: 1

A/Introns: 22/1

Best Local Similarity 54.5%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXXQTXLXNE 52
DB 1553 WLFQTLITNE 1563

RESULT 21
T10392
late expression factor 7 protein - *Orygia pseudotsugata* nuclear polyhedrosis virus
C:Species: *Orygia pseudotsugata* nuclear polyhedrosis virus, OpMNPV
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10392
R:Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F.
Virology 229, 381-399, 1997
A:Title: The sequence of the *Orygia pseudotsugata* multinucleocapsid nuclear polyhedrosis
A:Reference number: Z17011; MUID:97271300; PMID:9126251
A:Accession: T10392
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-211 <AHR>
A:Cross-references: EMBL:U75930; NID:g2934903; PIDN:AAC59122.1; PID:g1911369
C:Superfamily: *Orygia pseudotsugata* nuclear polyhedrosis virus late expression factor 7

Query Match 22.3%; Score 31; DB 2; Length 211;
Best Local Similarity 35.7%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXXXQTXLXNE 52
DB 133 WNWGLTRTLIHE 146

RESULT 22
AC0754
probable membrane protein STY2194 [imported] - *Salmonella enterica* subsp. *enterica* serov
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A:Note: this species has also been called *Salmonella typhi*
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0754
R:Parkhill, J.; Dougan, K.D.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AC0754
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-567 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05734.1; PID:g16503227; GSPDB:GN00176
C:Genetics:

Query Match 22.3%; Score 31; DB 2; Length 567;
Best Local Similarity 35.3%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXXXQTXLXNE 55
DB 393 WQAWHDPLRLYRGAL 409

RESULT 23
C64097
probable soluble lytic transglycosylase (BC 3.2.1.-) - *Haemophilus influenzae* (strain Rd
C:Species: *Haemophilus influenzae*
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 22-Jun-1999
C:Accession: C64097
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J

D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: C64097
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-593 <TIICR>
A:Cross-references: GB:U32765; GB:L42023; NID:g1573838; PIDN:AAC22487.1; PID:g1573843; TJ
C:Superfamily: soluble lytic transglycosylase
C:Keywords: glycosidase; hydrolase

Query Match 22.3%; Score 31; DB 2; Length 593;
Best Local Similarity 35.3%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXXXQTXLXNE 55
DB 300 WRYWEAKQDILLNTYK 316

RESULT 24
F91126
hypothetical protein ECs3982 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C:Species: *Escherichia coli*
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: F91126
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F91126
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-99 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA037405.1; PID:g13363455; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs3982

Query Match 21.6%; Score 30; DB 2; Length 99;
Best Local Similarity 28.6%; Pred. No. 22;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXXXQTXLXNE 52
DB 82 WSAWRLVKTLTKQ 95

RESULT 25
E85971
hypothetical protein Z4454 [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93
C:Species: *Escherichia coli*
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E85971
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-99 <STO>
A:Cross-references: GB:AE005174; NID:g12517690; PIDN:AGS58233.1; GSPDB:GN00145; UWGP:Z445
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z4454

Query Match 21.6%; Score 30; DB 2; Length 99;
Best Local Similarity 28.6%; Pred. No. 22;

```

Matches      4;  Conservative      2;  Mismatches      8;  Indels      0;  Gaps      0;

Qy      39  WXXWXXXQTXLXNE 52
      | | | | |
Db      82  WSAWRLVKTTLKQ 95

RESULT 26
A65099
hypothetical protein b3100 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
R/Accession: A65099
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
.A.; Rose, D.J.; Mau, B.; Shaoh, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: A65099
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-99 <BLAT>
A:Cross-references: GB:AF000392; GB:U000096; NID:g2367194; PIDN:AACT6135.1; PID:g1789487;
A:Experimental source: strain K-12, substrain MG1655

```

Query Match	21.6%	Score 30	DB 2	Length 99
Best Local Similarity	28.6%	Pred. No. 22		
Matches	4	Conservative	2	Mismatches 8
Indels				0
Gaps				0

RESULT 27
AE0895 conserved hypothetical protein STY3411 [imported] - *Salmonella enterica* subsp. *enterica*
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A:Note: this species has also been called *Salmonella typhi*
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AE0895
E:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AE0895
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-99 <PAR>
A:Cross-references: GB:AL513382; PID:CAD07755.1; PID:gl6504306; GSPDB:GN00176
C:Genetics:
A:Gene: STY3411

Query Match	21.6%	Score 30;	DB 2;	Length 99;
Best Local Similarity	28.6%;	Pred. No. 22;		
Matches	4;	Conservative	2;	Mismatches 8; Indels 0; Gaps 0;
QY	39	WXXXXXQTXLXNE	52	
	:	:	:	
Db	82	WSAARLVKTTLRQQ	95	

RESULT 28
C69026
acetylpolylamine aminohydrolase - Methanobacterium thermoautotrophicum (strain Delta H)
S.Species: Methanobacterium thermoautotrophicum
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C.Accession: C69026
R.Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997.
 A:Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: funct3
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: C69026
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-331 <MTH>
 A:Cross-references: GB:AE000887; GB:AE000666; NID:G2622289; PIDN:AA885683.1; PID:G2622302
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH194
 C:Superfamily: acetylpolymine aminohydrolase; RPD3/acuC homology
 F:3-298/Domain: RPD3/acuC homology <RAH1>

Query Match	21.6%	Score 30;	DB 1;	Length 331;
Best Local Similarity	45.5%	Pred. No. 68;		
Matches	5;	Conservative	0;	Mismatches 6; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXL 49
 | | | | |
 Db 263 WIGWFIHQTL 273
 | | | | |

RESULT 29
 T04525
 hypochelical protein Fl6Al6.150 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 04-Mar-2000
 C:Accession: T04525
 R:Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hoheisel, J.; Mewes
 submitted to the Protein Sequence Database, February 1999

RESULT 29

T04525
hypothetical protein F16A16.150 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 04-1-1999
C:Accession: T04525
R:Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.
submitted to the Protein Sequence Database, February 1999
A:Reference number: 215376
A:Accession: T04525
A:Molecule type: DNA
A:Residues: 1-331 <BEV>
A:Cross-references: EMBL:AL035353
A:Experimental source: cultivar Columbia; BAC clone F16A16
C:Genetics:
A:Map position: 4
A:Introns: 76/2; 269/2; 286/2
A:Note: F16A16.150

Query Match	21.6%	Score 30;	DB 2;	Length 331;
Best Local Similarity	38.5%	Pred. No. 68;		
Matches	5.	Conservative	0;	Mismatches 8;
				Indels 0;
				Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 267 WEKWLNEOKKLAN 27

RESULT 30

A86227
hypothetical protein [imported] - Arabidopsis thaliana

C_Species: Arabidopsis thaliana (mouse-ear cress)
C_Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C_Accession: A86227
R_Theologian: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, C.;
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.;
anssen, N.F.; Hughes, B.; Huizuar, L.
Nature 408, 816-820, 2000
A_Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A_Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A_Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A_Reference number: A86141; MUID:21016719; PMID:11130712

A;Reference number: A
A;Accession: A86227
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-385 <STO>
 A;Cross-references: GB:AE005172; NID:G3482913; PIDN:AAC33198.1; GSPDB:GN00141
 C;Genetics:
 A;Map position: 1

Query Match 21.6%; Score 30; DB 2; Length 385;
 Best Local Similarity 36.4%; Pred. No. 78;
 Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXTXNL 49
 DB 208 WAAWFIITQM 218

RESULT 31

T39230
 hypothetical protein SPAC9G1.07 - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T39230
 R;Churcher, C.M.; Gentles, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, August 1997
 A;Reference number: Z21837
 A;Accession: T39230
 A;Status: Preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-418 <CHU>
 A;Cross-references: EMBL:Z98763; PIDN:CAB11491.1; GSPDB:GN00066; SPDB:SPAC9G1.07
 A;Experimental source: strain 972h-; cosmid c9G1
 C;Genetics:
 A;Gene: SPDB:SPAC9G1.07
 A;Map position: 1

Query Match 21.6%; Score 30; DB 2; Length 418;
 Best Local Similarity 29.4%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 39 WXXWXXXTXNLNEXXL 55
 DB 319 WGSWSSQDTSKSSNL 335

RESULT 32

TYCHE2
 transcription factor ets-2 - chicken
 C;Species: Gallus gallus (chicken)
 C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C;Accession: S00386
 R;Boulukos, K.E.; Pognonec, P.; Begue, A.; Galibert, F.; Gesquiere, J.C.; Stehelin, D.;
 EMBO J 7, 697-705, 1988
 A;Title: Identification in chickens of an evolutionarily conserved cellular ets-2 gene
 A;Reference number: S00386; MUID:88283637; PMID:3293999
 A;Accession: S00386
 A;Molecule type: mRNA
 A;Residues: 1-479 <BCU>
 A;Cross-references: EMBL:X07202; NID:G63181; PIDN:CAA30178.1; PID:G63182
 A;Note: it is uncertain whether 1-Met, 9-Met, 18-Met, or 22-Met is the initiator codon
 C;Genetics:
 A;Gene: ets-2
 C;Superfamily: transcription factor ets; ets DNA-binding domain homology; ets RII regula
 C;Keywords: DNA binding; nucleus; phosphoprotein; proto-oncogene; transcription factor;
 F;94-167/Domain: ets RII regulatory region homology <ETR>
 F;134-167/Region: helix-loop-helix #status predicted
 F;375-453/Domain: ets DNA-binding domain homology
 F;415-421/Region: nuclear location signal

Query Match 21.6%; Score 30; DB 1; Length 479;
 Best Local Similarity 30.8%; Pred. No. 95;
 Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXTXNLN 51
 DB 117 WLAWATNEFSLAN 129

RESULT 33

C70606
 probable fadE34 protein - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C;Accession: C70606
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: C70606
 A;Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-711 <COL>
 A;Cross-references: GB:Z92774; GB:AL123456; NID:G3261729; PIDN:CAB07147.1; PID:e306713; I
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: fadE34

Query Match 21.6%; Score 30; DB 2; Length 711;
 Best Local Similarity 45.5%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXTXNLNE 52
 DB 571 WRLARTTILANE 581

RESULT 34

T29644
 hypothetical protein W01A11.3 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T29644
 R;Blanchard, M.; Bradshaw, H.
 submitted to the EMBL Data Library, July 1996
 A;Description: The sequence of C. elegans cosmid W01A11.
 A;Reference number: Z20658
 A;Accession: T29644
 A;Status: Preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-823 <BLA>
 A;Cross-references: EMBL:U64852; PIDN:AAB04966.1; GSPDB:GN00023; CESP:W01A11.3
 A;Experimental source: strain Bristol N2; clone W01A11
 C;Genetics:
 A;Gene: CESP:W01A11.3
 A;Map position: 5
 A;Introns: 15/3; 87/2; 137/2; 207/1; 258/3; 286/2; 419/3; 594/2; 729/2

Query Match 21.8%; Score 30; DB 2; Length 823;
 Best Local Similarity 30.8%; Pred. No. 1.6e+02;
 Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXTXNLN 51
 DB 641 WDAWNSQNDIRN 653

RESULT 35

B75573
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)
 C;Species: Deinococcus radiodurans
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C;Accession: B75573
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20035896; PMID:10567266
A:Accession: B75573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-546 <WHI>
A:Cross-references: GB:AB001863; GB:AB001825; NID:96460670; PIDN:AAF12422.1; PID:9646071
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0207
A:Map position: 2

Query Match 21.2%; Score 29.5; DB 2; Length 546;
Best Local Similarity 38.9%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 39 WXXWXXXQ-TLXNEXL 55
DB 281 WSAWRRQVTNLVRETAL 298

RESULT 36
S76185
hypothetical protein - Synecocystis sp. (strain PCC 6803)
A:Species: Synecocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S76185
R:Kaneko, T.; Sato, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76185
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-124 <KAN>
A:Cross-references: EMBL:D90914; GB:AB001339; NID:91653477; PIDN:BAAL8444.1; PID:d101917
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 20.9%; Score 29; DB 2; Length 124;
Best Local Similarity 28.6%; Pred. No. 45;
Matches 4; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQ-TLXNEXL 52
DB 97 WLSWLGHTNLHK 110

RESULT 37
JC5854
polyketide synthase (EC 2.3.1.41) chain 5 - Actinomadura hibisca
A:Species: Actinomadura hibisca
C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 17-Mar-1999
C:Accession: JC5854
R:Dairi, T.; Hamano, Y.; Igarashi, Y.; Furumai, T.; Oki, T.
BioSci. Biotechnol. Biochem. 61, 1445-1453, 1997
A:Title: Cloning and nucleotide sequence of the putative polyketide synthase genes for P
A:Reference number: JC5850; MUID:97480928; PMID:9339544
A:Accession: JC5854
A:Molecule type: DNA
A:Residues: 1-153 <DAI>
A:Cross-references: DDBJ:D87924
C:Comment: This enzyme catalyzes repeated condensation cycles of acyl-CoA, affecting the
C:Genetics:
A:Gene: pms5
C:Keywords: transferase

Query Match 20.9%; Score 29; DB 2; Length 153;
Best Local Similarity 36.4%; Pred. No. 54;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20035896; PMID:10567266
A:Accession: B75573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-546 <WHI>
A:Cross-references: GB:AB001863; GB:AB001825; NID:96460670; PIDN:AAF12422.1; PID:9646071
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0207
A:Map position: 2

Query Match 21.2%; Score 29.5; DB 2; Length 546;
Best Local Similarity 38.9%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 39 WXXWXXXQ-TLXNEXL 55
DB 281 WSAWRRQVTNLVRETAL 298

RESULT 36
S76185
hypothetical protein - Synecocystis sp. (strain PCC 6803)
A:Species: Synecocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S76185
R:Kaneko, T.; Sato, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76185
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-124 <KAN>
A:Cross-references: EMBL:D90914; GB:AB001339; NID:91653477; PIDN:BAAL8444.1; PID:d101917
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 20.9%; Score 29; DB 2; Length 124;
Best Local Similarity 28.6%; Pred. No. 45;
Matches 4; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQ-TLXNEXL 52
DB 97 WLSWLGHTNLHK 110

RESULT 37
JC5854
polyketide synthase (EC 2.3.1.41) chain 5 - Actinomadura hibisca
A:Species: Actinomadura hibisca
C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 17-Mar-1999
C:Accession: JC5854
R:Dairi, T.; Hamano, Y.; Igarashi, Y.; Furumai, T.; Oki, T.
BioSci. Biotechnol. Biochem. 61, 1445-1453, 1997
A:Title: Cloning and nucleotide sequence of the putative polyketide synthase genes for P
A:Reference number: JC5850; MUID:97480928; PMID:9339544
A:Accession: JC5854
A:Molecule type: DNA
A:Residues: 1-153 <DAI>
A:Cross-references: DDBJ:D87924
C:Comment: This enzyme catalyzes repeated condensation cycles of acyl-CoA, affecting the
C:Genetics:
A:Gene: pms5
C:Keywords: transferase

Query Match 20.9%; Score 29; DB 2; Length 153;
Best Local Similarity 36.4%; Pred. No. 54;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQ-TLXNEXL 49
DB 127 WAGWNAARTGL 137

RESULT 38
Q08CA1
cell division inhibitor sulA - Escherichia coli (strain K-12)
A:Species: Escherichia coli
C:Date: 24-Sep-1981 #sequence_revision 15-Oct-1996 #text_change 01-Mar-2002
C:Accession: A29016; A04468; I57720; E64836
R:Freund, R.; Braun, G.; Honore, N.; Cole, S.T.
Gene 52, 21-40, 1987
A:Title: Evolution of the enterobacterial sulA gene: a component of the SOS system encoded
A:Reference number: A29016; MUID:87248093; PMID:3297925
A:Accession: A29016
A:Molecule type: DNA
A:Residues: 1-169 <PRE>
R:Beck, E.; Bremer, E.
Nucleic Acids Res. 6, 3011-3024, 1980
A:Title: Nucleotide sequence of the gene ompA coding the outer membrane protein II of Es
A:Reference number: A93707; MUID:81053729; PMID:6253901
A:Accession: A04468
A:Molecule type: DNA
A:Residues: 1-144, 'IHSGKILSRHETTFRAKNSL', <BEC>
A:Experimental source: strain K-12
A:Note: this sequence has been revised by personal communication to the authors of refer
R:Cole, S.T.
Mol. Gen. Genet. 189, 400-404, 1983
A:Title: Characterisation of the promoter for the LexA regulated sulA gene of Escherichia
A:Reference number: I57720; MUID:83244178; PMID:6306396
A:Accession: I57720
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-10 <RES>
A:Cross-references: EMBL:V00358; NID:943023; PIDN:CAA23654.1; PID:943024
A:Note: the sequence represented here from this article is quoted from references A29016
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64836
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-169 <BLAT>
A:Cross-references: GB:AE000198; GB:U00096; NID:91787189; PIDN:AACT4044.1; PID:91787192;
A:Experimental source: strain K-12, substrain MGL655
C:Genetics:
A:Gene: sulA; sfiA
A:Map position: 22 min
C:Function:
A:Description: inducible cell division inhibitor; component of the SOS response; blocks f
x which may be the mechanism by which sulA inhibits cell division
A:Note: expression of sulA is repressed by lexA protein
C:Superfamily: cell division inhibitor sulA
C:Keywords: cell division control; SOS response

Query Match 20.9%; Score 29; DB 1; Length 169;
Best Local Similarity 35.7%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQ-TLXNEXL 52
DB 63 WQLWLTPOQKLISRE 76

RESULT 39
B29016
cell division inhibitor sulA - Salmonella typhimurium
A:Species: Salmonella typhimurium
C:Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 01-Feb-2002
C:Accession: B29016

R.Freudl, R.; Braun, G.; Honore, N.; Cole, S.T.
 Gene 52, 31-40, 1987
 A:Title: Evolution of the enterobacterial sula gene: a component of the SOS system encoded by the enterobacterial sula gene
 A:Reference number: A29016; MUID:87248093; PMID:3297925
 A:Accession: B29016
 A:Molecule type: DNA
 A:Residues: 1-169 <PRE>
 A:Cross-references: GB:M16324; NID:G154377; PIDN:AAA27230.1; PID:G154378
 C:Genetics:
 A:Gene: sula
 C:Function:
 A:Description: inducible cell division inhibitor
 A:Note: expression of sula is repressed by lexA protein
 C:Superfamily: cell division inhibitor sula
 C:Keywords: cell division control; SOS response

Query Match 20.9%; Score 29; DB 2; Length 169;
 Best Local Similarity 35.7%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
 Db 63 WQLWLTPOQKLSRE 76

RESULT 40
 C29016
 Cell division inhibitor sula - Enterobacter aerogenes
 C:Species: Enterobacter aerogenes
 C:Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 01-Feb-2002
 C:Accession: C29016
 R:Freudl, R.; Braun, G.; Honore, N.; Cole, S.T.
 Gene 52, 31-40, 1987
 A:Title: Evolution of the enterobacterial sula gene: a component of the SOS system encoded by the enterobacterial sula gene
 A:Reference number: A29016; MUID:87248093; PMID:3297925
 A:Accession: C29016
 A:Molecule type: DNA
 A:Residues: 1-169 <PRE>
 A:Cross-references: GB:M16467; NID:G148376; PIDN:AAA24810.1; PID:G148377
 C:Genetics:
 A:Gene: sula
 C:Function:
 A:Description: inducible cell division inhibitor
 A:Note: expression of sula is repressed by lexA protein
 C:Superfamily: cell division inhibitor sula
 C:Keywords: cell division control; SOS response

Query Match 20.9%; Score 29; DB 2; Length 169;
 Best Local Similarity 35.7%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
 Db 63 WQLWLTPOQKLSRE 76

RESULT 41
 B90759
 suppressor of lon [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: B90759
 R:Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 5, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic map of the enterohemorrhagic Escherichia coli O157:H7
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: B90759
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-169 <HAY>
 A:Cross-references: GB:BA00007; PIDN:BA034465.1; PID:G13360502; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:
 A:Gene: ECs1042
 C:Superfamily: cell division inhibitor sula

Query Match 20.9%; Score 29; DB 2; Length 169;
 Best Local Similarity 35.7%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
 Db 63 WQLWLTPOQKLSRE 76

RESULT 42
 H85622
 hypothetical protein sula [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85622
 R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, N.; Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85622
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-169 <STO>
 A:Cross-references: GB:AR005174; NID:G12514143; PIDN:AAG55444.1; GSPDB:GN00145; UWGP:Z133
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: sula
 C:Superfamily: cell division inhibitor sula

Query Match 20.9%; Score 29; DB 2; Length 169;
 Best Local Similarity 35.7%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
 Db 63 WQLWLTPOQKLSRE 76

RESULT 43
 AB0627
 cell division inhibitor [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain O157:H7, substrain EDL933)
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AB0627
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AB0627
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-169 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD08197.1; PID:G16502246; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY1092
 C:Superfamily: cell division inhibitor sula

Query Match 20.9%; Score 29; DB 2; Length 169;
 Best Local Similarity 35.7%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
 Db 63 WQLWLTPOQKLSRE 76


```
RESULT 44
AH0821
probable membrane protein STY2766 [imported] - Salmonella enterica subsp. enterica serov
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH0821
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Ouail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
C:Accession: AH0821
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02724.1; PID:g16503736; GSPDB:GN00176
C:Genetics:
A:Gene: STY2766
C:Superfamily: hypothetical protein HI0370

Query Match 20.9%; Score 29; DB 2; Length 206;
Best Local Similarity 44.4%; Pred. No. 71;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 39 WXXWXXXT 47
Db 39 WRYWSSHQT 47

RESULT 45
H75618
cob(I)alamin adenosyltransferase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2003
C:Accession: H75618
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Shen, M.; Vamathevan, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75618
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-207 <WHI>
A:Cross-references: GB:AE001826; NID:56460827; PIDN:AAF12570.1; PID:g6460866; TIGR:DRB00
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRB0008
A:Map position: megaplasmid
A:Genome: plasmid
A:Note: plasmid MPI
C:Superfamily: ATP:cob(I)alamin adenosyltransferase, CobA type

Query Match 20.9%; Score 29; DB 2; Length 207;
Best Local Similarity 29.4%; Pred. No. 71;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy 39 WXXWXXXTLXNEXXL 55
Db 148 WVAWEVEPVLKNRDL 164

RESULT 46
TI6103
hypothetical protein F19C7.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
```

```
C:Accession: TI6103
Rule, T.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid F19C7.
A:Reference number: Z18461
A:Accession: TI6103
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-310 <LET>
A:Cross-references: EMBL:U42439; NID:g1123106; PID:g1123109; PIDN:AAA83507.1; CESP:F19C7.
C:Genetics:
A:Gene: CESP:F19C7.3
A:Introns: 71/2; 178/3; 269/2
C:Superfamily: Caenorhabditis elegans hypothetical protein C24H12.11

Query Match 20.9%; Score 29; DB 2; Length 310;
Best Local Similarity 38.5%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXTLXN 51
Db 66 WIEWFYNTTKLON 78

RESULT 47
JC1069
nonstructure protein - rice dwarf virus
C:Species: rice dwarf virus
C>Date: 09-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 20-Sep-1999
C:Accession: JC1069
R:Chu, R.Y.; Zhang, X.; Pan, N.S.; Chen, Z.L.
Acta Bot. Sin. 35, 115-120, 1993
A>Title: The cDNA cloning and nucleotide sequence of the gene encoding nonstructure prote
A:Reference number: JC1069
A:Accession: JC1069
A:Molecule type: mRNA
A:Residues: 1-372 <CHU>
C:Genetics:
A:Map position: segment 10
C:Superfamily: wound tumor virus nonstructural protein Pns11

Query Match 20.9%; Score 29; DB 2; Length 372;
Best Local Similarity 23.5%; Pred. No. 1.2e+02;
Matches 4; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

Qy 39 WXXWXXXTLXNEXXL 55
Db 280 WRGWASTYAMFNQEQ 296

RESULT 48
AG2787
long-chain fatty acid transport protein fadL [imported] - Agrobacterium tumefaciens (str
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AG2787
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClella
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, F
ster, E.W.
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG2787
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-425 <KUR>
A:Cross-references: GB:AE008698; PIDN:AAL42717.1; PID:g17740155; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: fadL
A:Map position: circular chromosome
```

Search completed: September 30, 2004, 10:32:54
Job time : 18 secs

```
Query Match      20.9%; Score 29; DB 2; Length 425;
Best Local Similarity 44.4%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      39 WXXWXXXT 47
Db      282 WTDWSQIQ 290

RESULT 49
A97567
hypothetical protein AGR_C3154 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: A97567
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: A97567
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-429 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87490.1; PID:gi5156815; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C3154
A:Map position: circular chromosome

Query Match      20.9%; Score 29; DB 2; Length 429;
Best Local Similarity 44.4%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      39 WXXWXXXT 47
Db      286 WTDWSQIQ 294

RESULT 50
S76837
phosphopyruvate hydratase (EC 4.2.1.11) - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein slr0752
C:Species: Synechocystis sp.
A:Variety: FCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S76837
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76837
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-432 <KAN>
A:Cross-references: EMBL:D90917; GB:AB001339; NID:gi553836; PIDN:BAAL8749.1; PID:gi55383
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: eno
C:Superfamily: enolase
C:Keywords: carbon-oxygen lyase; gluconeogenesis; glycolysis; hydro-lyase

Query Match      20.9%; Score 29; DB 2; Length 432;
Best Local Similarity 29.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy      39 WXXWXXXTXLNEXXL 55
Db      297 WESWGLTSLGKTQL 313
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:28:51 ; Search time 9 Seconds
(without alignments)
376.062 Million cell updates/sec

Title: US-09-955-502-1

Perfect score: 139

Sequence: 1 MRRXXXCXXXXXXXXXXXXX.....QTXLXNEXLXXXXXXXXXX 65

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	35.3	90	1 Y953 YERPE	Qazhe7 yersinia pe
2	48	34.5	90	1 Y341 COXBU	Q3d06 xanthella bu
3	48	34.5	91	1 YP54 XANAC	Q3pjh7 xanthomonas
4	47	33.8	90	1 Y451 VIBCH	Q9kur4 vibrio chol
5	47	33.8	90	1 Y760 HAEIN	P44048 haemophilus
6	47	33.8	90	1 YD20 PASMU	Q9clb9 pasteurella
7	47	33.8	90	1 YF14 VIBVU	Q8dcs5 vibrio vuln
8	47	33.8	90	1 YD27 VIBPA	Q87li5 vibrio para
9	47	33.8	92	1 Y019 XANCP	Q9p823 xanthomonas
10	46	33.1	77	1 Y553 EUCAL	P57618 buchnera ap
11	46	33.1	88	1 Y419 NEINA	Q87lps neisseria m
12	46	33.1	90	1 YHE3 PSESM	Q87uf5 pseudomonas
13	46	33.1	92	1 YX69 SHEON	Q8ebx6 shewarella
14	45	32.4	90	1 Y285 PSBPK	Q88x49 pseudomonas
15	45	32.4	90	1 Y883 XYLFT	Q87d06 xylella fas
16	45	32.4	91	1 Y335 RALSO	Q8y010 ralstonia s
17	44	31.7	78	1 Y076 WIGBR	Q8dcs5 wiggiestwort
18	44	31.7	90	1 Y322 NITEU	Q82xf2 nitrosomonas
19	44	31.7	90	1 YFEB PSEAE	Q9th36 pseudomonas
20	44	31.7	90	1 YGXX ECOL6	Q8fe19 escherichia
21	44	31.7	90	1 YGXX ECOL1	P52065 escherichia
22	44	31.7	90	1 YGXX SALTY	Q8xfv6 salmonella
23	44	31.7	90	1 Y008 XYLPA	Q9pc73 xylella fas
24	43	30.9	78	1 Y535 BUCAP	Q8q925 buchnera ap
25	39	28.1	87	1 Y501 BUCBP	Q89a44 buchnera ap
26	34	24.5	589	1 PHBC ALCBU	P23608 a poly-beta
27	33	23.7	564	1 YEDQ ECO57	Q8xb92 escherichia
28	33	23.7	564	1 YEDQ ECOL1	P76330 escherichia
29	32	23.0	518	1 GSH1 BUCAP	P58994 buchnera ap
30	32	23.0	3655	1 YAMB SCFOP	Q10064 schizosacch
31	31	22.3	211	1 LEF7 NPVOP	O10362 orgyia pseu
32	31	22.3	567	1 YEDQ SALTI	Q8z5r0 salmonella
33	31	22.3	570	1 YEDQ SALTY	Q8znt5 salmonella

ALIGNMENTS

RESULT 1

Y953 YERPE STANDARD; PRT; 90 AA.

AC Q8ZHE7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0269 protein YPO0953/Y3340.
GN YPO0953 OR Y3340.

OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]

SEQUENCE FROM N.A.

RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth I., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN [2]

SEQUENCE FROM N.A.

RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;

RT "Genome sequence of Yersinia pestis KIM.";

RL J. Bacteriol. 184:4601-4611(2002).

CC -1- SIMILARITY: Belongs to the UPF0269 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; AJ414145; CAC89796.1; -

DR EMBL; AE013935; AAM86890.1; -

DR HAMAP; MF_00686; -; 1.

DR InterPro; IPR007457; DUF495.

P43888 haemophilus
Q47710 escherichia
O27262 methanobact
O14303 schizosacch
P10157 gallus gall
Q23064 caenorhabdi
P08846 escherichia
P08848 enterobacte
P08847 salmonella
Q85447 rice dwarf
Q85434 rice dwarf
P16594 rice dwarf
P87670 ebola virus
O11458 ebola virus
P60171 ebola virus
P60170 ebola virus
P77972 synchocyst

34 31 22.3 593 1 SLT_HABIN
35 30 21.6 99 1 YQVK_ECOLI
36 30 21.6 331 1 YB94_METTH
37 30 21.6 418 1 YB87_SCHPO
38 30 21.6 479 1 ETS2_CHICK
39 30 21.6 1041 1 UN83_CAEEL
40 29 20.9 169 1 SULA_ECOLI
41 29 20.9 169 1 SULA_ENTAE
42 29 20.9 169 1 SULA_SALTY
43 29 20.9 353 1 VP10_RDVA
44 29 20.9 353 1 VP10_RDVF
45 29 20.9 353 1 VP10_RDVO
46 29 20.9 364 1 VSGP_EBOEC
47 29 20.9 364 1 VSGP_EBOGA
48 29 20.9 364 1 VSGP_EBOZS
49 29 20.9 364 1 VSGP_EBOZM
50 29 20.9 432 1 ENO_SYNY3

RESULT 3


```
DR InterPro; IPR007457; DUF495.
DR Pfam; PF04362; DUF495; 1.
DR PIRSF; PIRSF029827; DUF495; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 92 AA; 10645 MW; F14506B181DB19E5 CRC64;

Query Match 33.8%; Score 47; DB 1; Length 92;
Best Local Similarity 23.6%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXXCXNXXXXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db 1 MSRTVFCVQCQDTEGLDFAPYBELGRIFAIQIKGAGQAWLAHQITMLINENL 55

RESULT 10
Y553_BUCAI
ID Y553_BUCAI STANDARD; PRT; 77 AA.
AC P57618;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0269 protein BU553.
GN BU553.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
sybiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -!- SIMILARITY: Belongs to the UPF0269 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001119; BAB13245.1; ALT_INIT.
CC HAMAP; MF_00686; -; 1.
CC InterPro; IPR007457; DUF495.
CC Pfam; PF04362; DUF495; 1.
CC PIRSF; PIRSF029827; DUF495; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 77 AA; 9511 MW; 00B049027CF480BF CRC64;

Query Match 33.1%; Score 46; DB 1; Length 77;
Best Local Similarity 23.6%; Pred. No. 0.0017;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXXCXNXXXXXXXXXXXXXPPXXXXXXXXXXXXXQTXLXNEXXL 55
Db 1 MRRIFCTFFKXSEGGQFQSGYFGKLGKIKYDQISKANEXKWKIEKQITLINENL 55

RESULT 11
Y419_NEIMA
ID Y419_NEIMA STANDARD; PRT; 88 AA.
AC Q9JQF5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0269 protein NMA0419/NMB2021.
GN NMA0419 OR NMB2021.
```

```
OS Neisseria meningitidis (serogroup A), and
OS Neisseria meningitidis (serogroup B)
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699; 491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajadream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis 22491.";
RL Nature 404:502-506(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
Eisen J.A., Ketchum K.A., Hood D.W., Feden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.B., DeBoy R., Peterson J.D., Hickey E.K.,
Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
Mason T., Ciecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B.,
RA Cotton M.D., Uitterback T.R., Khouri H., Qin H., Vamathevan J.,
Gill J., Searlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
RN [2]
RP SEQUENCE FROM N.A.
CC -!- SIMILARITY: Belongs to the UPF0269 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL162753; CAB83718.1; -.
CC EMBL; AE002552; AAF42344.1; -.
CC TIGR; NME2021; -.
CC HAMAP; MF_00686; -; 1.
CC InterPro; IPR007457; DUF495.
CC Pfam; PF04362; DUF495; 1.
CC PIRSF; PIRSF029827; DUF495; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 88 AA; 10180 MW; BEC6F2FBF097F2F1 CRC64;

Query Match 33.1%; Score 46; DB 1; Length 88;
Best Local Similarity 30.3%; Pred. No. 0.0019;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 23 PXXGXXXXXXXXXXXXWXXWXXQTXLXNEXXL 55
Db 23 PNEIGKRFENVSQEAQAAWTRHQTMLINENL 55

RESULT 12
YHE3_PSESM
ID YHE3_PSESM STANDARD; PRT; 90 AA.
AC Q87UF5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0269 protein PSPT05343.
GN PSPT05343.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
```


Query Match 32.4%; Score 45; DB 1; Length 90;
Best Local Similarity 30.3%; Pred. No. 0.0032;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 23 PXXGXXXXXXXQTXLNEXXL 55
DB 23 PGAKQDIFEHISOKAWADQKHQTLINEXRL 55

RESULT 15
Y883_XYLFT STANDARD; PRT; 90 AA.
AC Q87D06;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0269 protein PD0883.
GN PD0883.
OS Xylella fastidiosa (strain Temecula / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B., Moon D.H.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Teal S.M.,
RA Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA Marino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Bianco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa.";
RL J. Bacteriol. 185:1018-1026 (2003).
CC -!- SIMILARITY: Belongs to the UPF0269 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF012556; AA028748.1; -;
DR HAMAP; MF_00686; -; 1.
DR InterPro; IPR007457; DUF495.
DR Pfam; PF04362; DUF495; 1.
DR PIRSF; PIRSF039827; DUF495; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 90 AA; 10430 MW; B5C6936496F4171D CRC64;

Query Match 32.4%; Score 45; DB 1; Length 90;
Best Local Similarity 47.1%; Pred. No. 0.0032;
Matches 8; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLNEXXL 55
DB 39 WAAWLAHQTLINEXRL 55

RESULT 16
YC35_RALSO STANDARD; PRT; 91 AA.
ID YC35_RALSO
AC Q8Y010;
DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0269 protein RSC1235.
GN RSC1235 OR RS02742.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RA MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Attiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chaudler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguiet P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502 (2002).
CC -!- SIMILARITY: Belongs to the UPF0269 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL646063; CAD14937.1; -;
DR HAMAP; MF_00686; -; 1.
DR InterPro; IPR007457; DUF495.
DR Pfam; PF04362; DUF495; 1.
DR PIRSF; PIRSF039827; DUF495; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 91 AA; 10321 MW; BA7PAD0032CB3C9B CRC64;

Query Match 32.4%; Score 45; DB 1; Length 91;
Best Local Similarity 26.8%; Pred. No. 0.0032;
Matches 11; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 23 PXXGXXXXXXXQTXLNEXXLXXXXX 63
DB 23 PGBLKKIWSVSKEAWAGLKHQTLINEXRLNADTRAR 63

RESULT 17
Y076_WIGBR STANDARD; PRT; 78 AA.
ID Y076_WIGBR
AC Q8D3C5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0269 protein WIGBR0760.
GN WIGBR0760.
OS Wigglesworthia glossinidia brevipalpis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Wigglesworthia.
OX NCBI_TaxID=36870;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22377718; PubMed=12219091;
RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,
RA Aksoy S.;
RT "Genome sequence of the endocellular obligate symbiont of tsetse
RT flies, Wigglesworthia glossinidia.";
RL Nat. Genet. 32:402-407 (2002).
CC -!- SIMILARITY: Belongs to the UPF0269 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

RT of uropathogenic *Escherichia coli*.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).

CC -1- SIMILARITY: Belongs to the UPF0269 family.

CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

DR ENBL; A5016766; AAN81998.1; -

DR HAWAP; MF_00686; -; 1.

DR InterPro; IPR007457; DUF495.

DR Pfam; PF04362; DUF495; 1.

DR Complete proteome.

FT INIT MET 0 0 BY SIMILARITY.

FT SEQUENCE 90 AA; 10805 MW; D7C66C2A35B563692 CRC64;

CC -----

Query Match 31.7%; Score 44; DB 1; Length 90;

Best Local Similarity 30.3%; Pred. No. 0.0053;

Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 23 PXXGXXXXXXXWXXWXXQXXKXNEXL 55

DB 22 PGLGLRYNISKAWAQWQHQTWLINEKX 54

RESULT 21

YGGX_ECOLI STANDARD; PRT; 90 AA.

ID YGGX_ECOLI STANDARD; PRT; 90 AA.

AC P22085;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE UPF0269 protein yggX.

DE NGX OR B2962 OR Q24307 OR ECS3838 OR SF2959 OR S3162.

OS *Escherichia coli*.

OS *Escherichia coli* O157:H7, and

OS *Shigella flexneri*.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; *Escherichia*.

OX NCBI_TaxID=562, 83334, 623;

[1]

SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Grainger J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.

RT "The complete genome sequence of *Escherichia coli* K-12.";

RL Science 277:1453-1474 (1997).

[2]

SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;

RX MEDLINE=21074935; PubMed=11206551;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,

RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,

RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

RA Welch R.A., Blattner F.R.

RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";

RL Nature 409:529-533 (2001).

[3]

SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;

RX MEDLINE=21156231; PubMed=11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,

RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,

RA Tida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,

RA InterPro; IPR007457; DUF495.

RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;

RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*

RT O157:H7 and genomic comparison with a laboratory strain K-12.";

RL Nucleic Acids Res. 28:11-22 (2001).

[4]

RP PARTIAL SEQUENCE OF 1-12.

RC SPECIES=E.coli; STRAIN=K12 / EMG2;

RX MEDLINE=97443975; PubMed=9298646;

RA Link A.J., Robinson K., Church G.M.;

RT "Comparing the predicted and observed properties of proteins encoded

RT in the genome of *Escherichia coli* K-12.";

RL Electrophoresis 18:1259-1313 (1997).

[5]

RP SEQUENCE OF 1-10.

RC SPECIES=E.coli; STRAIN=K12;

RX MEDLINE=99085675; PubMed=9868784;

RA Wasinger V.C., Humphrey-Smith I.;

RT "Small genes/gene-products in *Escherichia coli* K-12.";

RL FEMS Microbiol. Lett. 169:375-382 (1998).

[6]

RP SEQUENCE FROM N.A.

RC SPECIES=S. flexneri; STRAIN=301 / Serotype 2a;

RX MEDLINE=22272406; PubMed=12384590;

RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,

RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,

RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,

RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,

RA Yu J.

RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity

RT through comparison with genomes of *Escherichia coli* K12 and O157.";

RL Nucleic Acids Res. 30:4432-4441 (2002).

[7]

RP SEQUENCE FROM N.A.

RC SPECIES=S. flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;

RX MEDLINE=22590274; PubMed=12704152;

RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,

RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,

RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,

RA Schwartz D.C., Blattner F.R.;

RT "Complete genome sequence and comparative genomics of *Shigella*

RT flexneri serotype 2a strain 2457T.";

RL Infect. Immun. 71:2775-2786 (2003).

[8]

RP IDENTIFICATION BY MASS SPECTROMETRY.

RC SPECIES=E.coli;

RX MEDLINE=99420866; PubMed=10493123;

RA Fountoulakis M., Takacs M.-F., Berndt P., Langen H., Takacs B.;

RT "Enrichment of low abundance proteins of *Escherichia coli* by

RT hydroxyapatite chromatography.";

RL Electrophoresis 20:2181-2195 (1999).

CC -1- SIMILARITY: Belongs to the UPF0269 family.

CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

DR ENBL; U28377; AAA69129.1; -

DR ENBL; AE000378; AAC75999.1; -

DR ENBL; AE005526; AAG58093.1; -

DR ENBL; AP002563; BAB37261.1; -

DR ENBL; AE015308; AAN44440.1; -

DR ENBL; AE016988; AAP18264.1; -

DR PIR; A65082; A65082.

DR PIR; A85954; A85954.

DR PIR; F91108; F91108.

DR SWISS-2DPAGE; P52065; COLI.

DR EcoGene; EG12984; yggX.

DR HAWAP; MF_00686; -; 1.

DR InterPro; IPR007457; DUF495.

```

DR PFAM: PF04362; DUF495; 1.
DR PIRSF: PIRSF029827; DUF495; 1.
KW Complete proteome.
FT INIT_MET 0
SQ SEQUENCE 90 AA; 10821 MW; D7C66C2A35E62402 CRC64;

Query Match 31.7%; Score 44; DB 1; Length 90;
Best Local Similarity 30.3%; Pred. No. 0.0053;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 23 PXXGXXXXXXXWXXWXXWXXQTXLXNXXL 55
DB 22 PGLGKRIYNEISKEAWAQWQHKTMLINEK 54

RESULT 22
YGGX_SALTY STANDARD; PRT; 90 AA.
AC Q8XFV6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UPF0269 protein YGGX.
GN YGGX OR STM3111 OR STY3266 OR T3024.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Churcher C., Mungall K.L., Bentley K., Chillingworth T., Connor P.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K., Moulton S., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."
RL J. Bacteriol. 185:2330-2337 (2003).
CC -!- SIMILARITY: Belongs to the UPF0269 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: AE008843; AAL21986.1; -
CC DR EMBL: AL627277; CAD02936.1; -
CC DR EMBL: AE016844; AA070576.1; -
CC DR StyGene; SG?????; YGGX.
CC DR HAMAP; MF_00686; -; 1.
CC DR InterPro; IPR007457; DUF495.
CC DR Pfam; PF04362; DUF495; 1.
CC DR PIRSF; PIRSF029827; DUF495; 1.
CC KW Complete proteome.
CC FT INIT_MET 0 BY SIMILARITY.
CC SQ SEQUENCE 90 AA; 10768 MW; B3B6A1EE2255E3C CRC64;

Query Match 31.7%; Score 44; DB 1; Length 90;
Best Local Similarity 30.3%; Pred. No. 0.0053;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 23 PXXGXXXXXXXWXXWXXWXXQTXLXNXXL 55
DB 22 PGLGKRIYNEISKEAWAQWQHKTMLINEK 54

RESULT 23
XY08_XYLFA STANDARD; PRT; 90 AA.
ID XY08_XYLFA
AC Q9PC73;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0269 protein Xf1908.
GN Xf1908.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=985C.
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barrios M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.N.B.N., Madeira H.M.F., Marino C.U.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Weidman J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159 (2000).
CC -!- SIMILARITY: Belongs to the UPF0269 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004010; AAF84714.1; ALT_INIT.
DR HAMAP; MF 00686; -; 1.
DR InterPro; IPR007457; DUF495.
DR Pfam; PF04362; DUF495; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 90 AA; 10555 MW; E5BE83BCA6D104A3 CRC64;

Query Match      31.7%; Score 44; DB 1; Length 90;
Best Local Similarity 47.1%; Pred. No. 0.0053;
Matches 8; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXL 55
DB 39 WAAWLVTQTLNENRL 55

RESULT 24
Y535_BUCAP STANDARD; PRT; 78 AA.
AC Q8K925;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0269 protein BUG535.
GN BUG535.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
SQ SEQUENCE FROM N.A.
MEDLINE=22084549; PubMed=12089438;
RA Tamas I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379(2002).
CC -1- SIMILARITY: Belongs to the UPF0269 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014127; AAW68076.1; -.
DR HAMAP; MF 00686; -; 1.
DR InterPro; IPR007457; DUF495.
DR Pfam; PF04362; DUF495; 1.
DR PIRSF; PIRSF029827; DUF495; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 78 AA; 9564 MW; AB57086986777F44 CRC64;

Query Match      30.9%; Score 43; DB 1; Length 78;
Best Local Similarity 30.3%; Pred. NO. 0.0078;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 23 PXXKGGXXXXXXKXXXXXXQTXLXNEXL 55
DB 23 PGKLGKYNQISKIAKWKWLOQTLLNEEKL 55

RESULT 25
Y501_BUCBP STANDARD; PRT; 87 AA.
ID Y501_BUCBP
AC Q89A48;
DT 10-OCT-2003 (Rel. 42, Created)

```

```

DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0269 protein bfp501.
GN BFP501.
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135842;
RN [1]
SQ SEQUENCE FROM N.A.
MEDLINE=22426901; PubMed=12522265;
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT "Reductive genome evolution in Buchnera aphidicola.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -1- SIMILARITY: Belongs to the UPF0269 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014017; AAO27206.1; -.
DR HAMAP; MF 00686; -; 1.
DR InterPro; IPR007457; DUF495.
DR Pfam; PF04362; DUF495; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 87 AA; 10676 MW; C7270142539D1223 CRC64;

Query Match      28.1%; Score 39; DB 1; Length 87;
Best Local Similarity 26.7%; Pred. No. 0.065;
Matches 8; Conservative 1; Mismatches 21; Indels 0; Gaps 0;

QY 23 PXXKGGXXXXXXKXXXXXXQTXLXNE 52
DB 29 PGLGLGHKYNEISKLANWKWLQQTIIINE 58

RESULT 26
PHBC_ALCEU STANDARD; PRT; 589 AA.
ID PHBC_ALCEU
AC P23608;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Poly-beta-hydroxybutyrate polymerase (EC 2.3.1.-) (Poly(3-
DE hydroxybutyrate) polymerase) (PHB polymerase) (PHA synthase)
DE hydroxyalkanoate polymerase) (PHA polymerase) (PHA synthase)
GN PHBC.
OS Alcaligenes eutrophus (Ralstonia eutropha).
CC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
CC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=510;
RN [1]
SQ SEQUENCE FROM N.A.
STRAIN=H16 / DSM 428 / ATCC 17699;
MEDLINE=89359357; PubMed=2670936;
RA Peoples O.P., Sinskey A.J.;
RT "Poly-beta-hydroxybutyrate (PHB) biosynthesis in Alcaligenes
RT eutrophus H16. Identification and characterization of the PHB
RT polymerase gene (phbC).";
RL J. Biol. Chem. 264:15298-15303(1989).
RN [2]
SQ SEQUENCE OF 1-219 FROM N.A.
MEDLINE=91100279; PubMed=1987116;
RA Schubert P., Krueger N., Steinbuechel A.;
RT "Molecular analysis of the Alcaligenes eutrophus
RT poly(3-hydroxybutyrate) biosynthetic operon: identification of the N

```

terminus of poly(3-hydroxybutyrate) synthase and identification of the promoter.";

RT J. Bacteriol. 173:168-175(1991).

RL -!- FUNCTION: POLYMERIZES D(-)-3-HYDROXYBUTYRYL-COA TO CREATE PHB

CC WHICH CONSISTS OF THOUSANDS OF HYDROXYBUTYRATE MOLECULES LINKED

CC END TO END. PHB SERVES AS AN INTRACELLULAR ENERGY RESERVE MATERIAL

CC WHEN CELLS GROW UNDER CONDITIONS OF NUTRIENT LIMITATION.

CC -!- PATHWAY: Poly-beta-hydroxybutyrate biosynthesis; third step.

CC -!- SUBUNIT: Monomer (Probable).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- BIOTECHNOLOGY: Forms a biodegradable plastic that is degraded

CC naturally and completely by bacteria into carbon dioxide and

CC water. Utilized in the medical industry, plates made from PHA-

CC based plastics can be left in place to help heal fractured bones.

CC After the bone has healed, the plastic slowly breaks down in the

CC body. Utilized by Imperial Chemical Industries (ICI) to produce a

CC PHB-PHV (poly-B-valerate) copolymer solid under the trade name

CC 'Biopol'. Biopol is used as packaging material. The PHB-PHV

CC copolymer consists of approximately 20% PHV and 80% PHB. It can be

CC synthesized by incorporating glucose and valeric acid into the

CC medium. PHB-PHV is stronger and more flexible than regular PHB.

CC Under industrial conditions, 80% or higher of the cell dry weight

CC of A. eutrophus usually consists of the PHB-PHV copolymer.

CC -!- SIMILARITY: Belongs to the PHA/PHB synthase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; J05003; AAA21975.1; -.

DR EMBL; M64341; AAA21979.1; -.

DR PIR; A34341; A34341.

DR InterPro; IPR00073; A/b hydrolase.

DR Pfam; PF00561; abhydrolase; 1.

DR PHB biosynthesis; Transferase; Acyltransferase.

FT ACT_SITE 319 319 POTENTIAL.

FT ACT_SITE 319 319 POTENTIAL.

SQ SEQUENCE 589 AA; 64316 MW; A822F35CF70D8B68 CRC64;

Query Match 24.5%; Score 34; DB 1; Length 589;

Best Local Similarity 42.9%; Pred. No. 4.7;

Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52

DB 436 WYCWYLRHTYLQNE 449

RESULT 27

ID YEDQ ECO57 STANDARD; PRT; 564 AA.

AC Q8XB92.

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein yedQ.

GN YEDQ OR Z3047 OR ECS2694.

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=83334;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / EDL933 / ATCC 700927;

RX MEDLINE=21074935; PubMed=11206553;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,

RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,

RA Postfai G., Hackett J., Xlink S., Boutin A., Shao Y., Miller L.,

RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

Welch R.A., Blattner F.R.,

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";

RL Nature 409:529-533(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / RIMD 0509952;

RX MEDLINE=21156231; PubMed=11258796;

RA Hayashi I., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,

RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,

RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,

RA Kuhara S., Shiba T., Hattori M., Shinagawa H.

RT "Complete genome sequence of enterohemorrhagic Escherichia coli

RT O157:H7 and genomic comparison with a laboratory strain K-12.";

RL DNA Res. 8:11-22(2001).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -!- SIMILARITY: BELONGS TO THE YAC / YFIN (E.COLI), YHCK (B.SUBTILIS)

CC FAMILY.

CC -!- SIMILARITY: Contains 1 GDEF domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE005417; AAG56970.1; -.

DR EMBL; AP002559; BAB36117.1; -.

DR PIR; F90965; F90965.

DR InterPro; IPR00160; GDEF.

DR Pfam; PF00990; GDEF; 1.

DR SMART; SM00267; DUF1; 1.

DR TIGRFAMs; TIGR00254; GDEF; 1.

DR PROSITE; PS00887; GDEF; 1.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 20 40 POTENTIAL.

FT TRANSMEM 360 380 POTENTIAL.

FT DOMAIN 428 563 GDEF.

SQ SEQUENCE 564 AA; 64287 MW; BEBC2286ADBACEB0 CRC64;

Query Match 23.7%; Score 33; DB 1; Length 564;

Best Local Similarity 35.3%; Pred. No. 7.4;

Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXI 55

DB 396 WQAWHDTLRLYRGAL 412

RESULT 28

ID YEDQ ECOLI STANDARD; PRT; 564 AA.

AC P76330; P94746;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein yedQ.

GN YEDQ OR B1956.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MGL655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

MAU B., Shao Y.

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474 (1997).

```

RN RP
RC SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitagawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Sivasubramaniam S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map.";
PL DNA Res. 3:379-392(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE YABC / YFYN (E.COLI), YHCK (B.SUBTILIS)
CC FAMILY.
CC -1- SIMILARITY: Contains 1 GDEF domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000287; AAC75022.1; ALT_INIT.
CC EMBL; D90835; BAA15784.1; -.
CC EcoGene; EGI4040; yedQ.
CC InterPro; IPR000160; GDEF.
CC Pfam; PF00990; GDEF; 1.
CC SMART; SM00267; DUF1; 1.
CC TIGRFAMs; TIGR00254; GDEF; 1.
CC PROSITE; PS00887; GDEF; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 20 40 POTENTIAL.
FT DOMAIN 360 380 POTENTIAL.
FT DOMAIN 428 563 GDEF.
SQ SEQUENCE 564 AA; 64283 MW; 05FB02C1BE2A8938 CRC64;

Query Match 23.7%; Score 33; DB 1; Length 564;
Best Local Similarity 35.3%; Pred. No. 7.4; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 11;

QY 39 WXXWXXXTQLXNEXXL 55
DB 396 WQAWHDTITLYNRCAL 412

RESULT 29
GSH1_BUCAP
AC P58394;
DC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamate--cysteine ligase (EC 6.3.2.2) (Gamma-glutamylcysteine
DE synthetase) (GCS).
GN SYNTH OR BUSG392.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamas I., Klason L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstrom J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379(2002).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + L-cysteine = ADP +
CC phosphate + gamma-L-glutamyl-L-cysteine.
CC -1- PATHWAY: Glutathione biosynthesis; first step.

```

```

CC -1- SIMILARITY: Belongs to the glutamate--cysteine ligase family 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014115; AAM67944.1; -.
CC HAMAP; MF_00578; -; 1.
CC InterPro; IPR007370; Glu_cys_ligase.
CC InterPro; IPR006334; Glu_cys_ligase.
CC Pfam; PF04262; glu_cys_ligase; 1.
CC TIGRFAMs; TIGR01434; glu_cys_ligase; 1.
KW Glutathione biosynthesis; ligase; Complete proteome.
SQ SEQUENCE 518 AA; 61250 MW; 53E76B05654807F3 CRC64;

Query Match 23.0%; Score 32; DB 1; Length 518;
Best Local Similarity 35.7%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXTQLXNE 52
DB 160 WKNWKKSEINLXNE 173

RESULT 30
YAME_SCHPO
ID YAME_SCHPO STANDARD; PRT; 3655 AA.
AC Q10084;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hypothetical protein ClF5.11c in chromosome I.
GN SPAC1F5.11C
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RX Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Art R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002)
CC -1- SIMILARITY: STRONG, TO YEAST YHR099W.

```


Query Match 22.3%; Score 31; DB 1; Length 567;
 Best Local Similarity 35.3%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLNEXXL 55
 DB 393 QAWHDFLTRLYNRGAL 409

RESULT 33
 YEDQ_SALTY STANDARD; PRT; 570 AA.
 AC Q8ZNTS;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein yedQ.
 GN YEDQ OR STM1987.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856 (2001).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: BELONGS TO THE YALC / YFIN (E.COLI), YHCK (B.SUBTILIS)
 CC FAMILY.
 CC -!- SIMILARITY: Contains 1 GGDEF domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE008788; AAL20897.1; --
 DR StyGene; SG????; yedQ.
 DR InterPro; IPR000160; GGDEF.
 DR Pfam; PF00990; GGDEF; 1.
 DR SMART; SM00267; DUF1; 1.
 DR TIGRfam; TIGR00254; GGDEF; 1.
 DR PROSITE; PS50887; GGDEF; 1.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 20 40 POTENTIAL.
 FT TRANSMEM 360 380 POTENTIAL.
 FT DOMAIN 428 563 GGDEF.
 FT SEQUENCE 570 AA; 65429 MW; BF8FDF0CF8949925 CRC64;
 SQ

Query Match 22.3%; Score 31; DB 1; Length 570;
 Best Local Similarity 35.3%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLNEXXL 55
 DB 396 QAWHDFLTRLYNRGAL 412

RESULT 34
 SLT_HAEIN STANDARD; PRT; 593 AA.
 ID SLT_HAEIN PRT; 593 AA.
 AC P44588;

DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Putative soluble lytic murein transglycosylase precursor (EC 3.2.1.-).
 GN SLT OR HI0829.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA McEwen R.A., Tomb J.F., Dougherty B.A., Merrick J.M.,
 RA Klenzmann K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512 (1995).
 CC -!- FUNCTION: MUREIN-DEGRADING ENZYME. MAY PLAY A ROLE IN RECYCLING
 CC OF MUROPETIDES DURING CELL ELONGATION AND/OR CELL DIVISION
 CC (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: Cleavage of the beta-1,4-glycosidic bond
 CC between N-acetylmuramic acid and N-acetylglucosamine residues,
 CC thereby conserving the energy in a newly synthesized
 CC 1,6-anhydrobond in the muremic acid residue.
 CC -!- SUBCELLULAR LOCATION: Periplasmic (Potential).
 CC -!- SIMILARITY: Belongs to the transglycosylase slt family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U32765; AAC22487.1; --
 DR PIR; C64097; C64097.
 DR HSPR; P03810; IQSA.
 DR TIGR; HI0829; --
 DR InterPro; IPR008939; Muramidase_bact.
 DR InterPro; IPR008258; SLT dom.
 DR InterPro; IPR000189; Transglyc_AS.
 DR Pfam; PF01464; SLT; 1.
 DR PROSITE; PS00922; TRANSGLYCOSYLASE; 1.
 KW Cell wall; Hydrolase; Glycosidase; Periplasmic; Signal;
 KW Complete proteome.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 593 PUTATIVE SOLUBLE LYTIC MUREIN
 FT DOMAIN 440 529 TRANSGLYCOSYLASE.
 FT ACT SITE 453 453 SLT-TYPE DOMAIN.
 FT SEQUENCE 593 AA; 68691 MW; 774DD3D38217CEE CRC64;
 SQ

Query Match 22.3%; Score 31; DB 1; Length 593;
 Best Local Similarity 35.3%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLNEXXL 55
 DB 300 WRYWEAKQDILKNTKKL 316

RESULT 35
 YQK_ECOLI STANDARD; PRT; 99 AA.
 ID YQK_ECOLI PRT; 99 AA.

Q47710;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein YqJK.
 GN YQJK OR B3100.
 OS Escherichia coli.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U18997; AAA57904.1; -;
 CC EXBL: AE000392; AAC76135.1; -;
 CC PIR: A65099; A65099.
 CC Ecogene; EGI4314; YqJK.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 99 AA; 11811 MW; DAD56718FFAF4086 CRC64;

 Query Match 21.6%; Score 30; DB 1; Length 99;
 Best Local Similarity 28.6%; Pred. No. 7;
 Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 39 WXXWXXXQTXLXNE 52
 DB 82 WSAWRLVKTKLQK 95

 RESULT 36
 YB94_METH
 ID YB94_METH STANDARD; PRT; 331 AA.
 AC O27262;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein MTH1194.
 GN MTH1194.
 OS Methanobacterium thermoautotrophicum.
 CC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 CC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Decuette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lum W., Potier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwni N., Caruso A., Bush D., Safer H., Patwell D., Prabakar S.,
 RA Mcbougall S., Shmer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -----
 CC -!- FUNCTION: Putative deacetylase (By similarity).
 CC -----

CC -!- SIMILARITY: Belongs to the histone deacetylase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000887; AAB85683.1; -;
 CC PIR: C69026; C69026.
 CC InterPro: IPR000286; His deacetylase.
 CC Pfam: PF00850; Hist deacetyl; 1.
 CC PRINTS: PRO1270; HDASUPER.
 KW Hypothetical protein; Hydrolase; Complete proteome.
 SQ SEQUENCE 331 AA; 36722 MW; 489054F32965EDCF CRC64;

 Query Match 21.6%; Score 30; DB 1; Length 331;
 Best Local Similarity 45.5%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 39 WXXWXXXQTXL 49
 DB 263 WIGWFIHQTL 273

 RESULT 37
 YB87_SCHPO
 ID YB87_SCHPO STANDARD; PRT; 418 AA.
 AC O14303;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C9G1.07 in chromosome I.
 GN SPAC9G1.07.
 OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstraelen E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehtach H., Reinhardt R., Pohl T.M.,
 RA Ezer P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Mottier S.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomycetes pombe.";
 RL Nature 415:871-880(2002).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF338767; AAL15621.1; -;
 DR EMBL; U64852; AAB04966.2; -;
 DR WormPep; W01A11.3; CE31077.

KW Nuclear protein; Transmembrane; Developmental protein; Coiled coil;
 KW Alternative splicing.

FT TRANSMEM 1005 1024

FT ANCHOR FOR TYPE IV MEMBRANE PROTEIN

FT (POTENTIAL).

FT COILED COIL (POTENTIAL).

FT COILED COIL (POTENTIAL).

FT Missing (in isoform b).

FT /FTid=VSP_007083.

FT Missing (in isoform c).

FT /FTid=VSP_007084.

FT SEQUENCE 1041 AA; 117821 MW; BB403A9A7C41A01F CRC64;

Query Match 21.6%; Score 30; DB 1; Length 1041;

Best Local Similarity 30.8%; Pred. No. 59;

Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXTXN 51

Db 859 WDANRQNDIRN 871

RESULT 40

SULA_ECOLI

ID SULA_ECOLI STANDARD; PRT; 169 AA.

AC P08846; F03840; P71224;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cell division inhibitor.

GN SULA OR SP1A OR B0958 OR Z1308 OR ECS1042.

OS Escherichia coli, and

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562; 83334;

[1]

SEQUENCE FROM N.A.

RX MEDLINE=87248093; PubMed=3297925;

RA Freud R., Braun G., Honore N., Cole S.T.;

RT "Evolution of the enterobacterial sula gene: a component of the SOS

RT system encoding an inhibitor of cell division.";

RL Gene 52.31-40(1987).

[2]

SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=81053729; PubMed=62533901;

RA Beck E., Bremer E.;

RT "Nucleotide sequence of the gene ompA coding the outer membrane

RT protein II of Escherichia coli K-12.";

RL Nucleic Acids Res. 8:3011-3024 (1980).

[3]

SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474 (1997).

[4]

SEQUENCE FROM N.A.

RP STRAIN=K12;

RX

RA

RA

RA

RA

RA

RA

RA

RA

RT

RT

RL

RL

RN

RN

RP

RP

RC

RC

RA

RA

RA

RA

RA

RA

RA

RA

RA

RT

RT

RL

RL

RN

RN

RP

RP

RX

RA

RA

RT

RT

RL

RL

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

MEDLINE=97061202; PubMed=8905232;

RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,

RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,

RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saio N.,

RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,

RA Yano M., Horiuchi T.;

RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome

RT corresponding to the 12.7-28.0 min region on the linkage map.";

RL DNA Res. 3:137-155 (1996).

[5]

SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / EDL933 / ATCC 700927;

RC MEDLINE=21074935; PubMed=11208551;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,

RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,

RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

RA Welch R.A., Blattner F.R.;

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";

RL Nature 409:529-533 (2001).

[6]

SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / RIMD 0509952;

RC MEDLINE=21156231; PubMed=11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,

RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,

RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,

RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;

RT "Complete genome sequence of enterohemorrhagic Escherichia coli

RT O157:H7 and genomic comparison with a laboratory strain K-12.";

RL DNA Res. 8:11-22 (2001).

[7]

SEQUENCE OF 1-9 FROM N.A.

RC MEDLINE=83244178; PubMed=6306396;

RA Cole S.T.;

RT "Characterisation of the promoter for the LexA regulated sula gene of

RT Escherichia coli.";

RL Mol. Gen. Genet. 189:400-404 (1983).

CC -!- FUNCTION: SULA IS A COMPONENT OF THE SOS SYSTEM AND AN INHIBITOR

CC OF CELL DIVISION. ACCUMULATION OF SULA CAUSES RAPID CESSATION OF

CC CELL DIVISION AND THE APPEARANCE OF LONG, NON-SEPTATE FILAMENTS.

CC THE EXPRESSION OF SULA IS RERESSED BY LEXA PROTEIN. FTSZ SEEMS

CC TO BE THE TARGET OF SULA.

CC -!- SUBCELLULAR LOCATION: Inner membrane.

CC -!- MISCELLANEOUS: THE N-TERMINUS OF SULA PROTEIN MIGHT BE INVOLVED IN

CC RECOGNIZING THE CELL DIVISION APPARATUS.

CC -!- SIMILARITY: THE C-TERMINUS SHOWS SIMILARITY TO THE N PROTEIN OF

CC PHAGE LAMBDA.

CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 145

CC ONWARD DUE TO A FRAMESHIFT.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

EMBL; V00307; CAA23587.1; ALT_FRAME.

EMBL; V00358; CAA23654.1; -;

EMBL; AE000198; AAC74044.1; -;

EMBL; D90733; BAA35716.1; -;

EMBL; D90734; BAA35723.1; -;

EMBL; AE005286; AAG55444.1; -;

EMBL; AP002554; BAB34465.1; -;

EMBL; A29016; QOECAL.

PIR; A29016; B30759.

PIR; B30759; B30759.

EcoGene; EGI0984; sula.

InterPro; IPR004596; Sula.

Pfam; PF03846; Sula; 1.

```
DR TIGRFAMs; TIGR00623; sula: 1.
KW Cell division; Septation; SOS response; Inner membrane;
KW Complete proteome.
FT DOMAIN 31 149 CONSERVED REGION.
FT DOMAIN 160 169 LON PROTEIN BINDING SITE (PROBABLE).
SQ SEQUENCE 169 AA; 18801 MW; C76B4493773C77C2 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 169;
Best Local Similarity 35.7%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 63 WQLWLTPQOKLSRE 76

RESULT 41
SULA_ENTAE
ID SULA_ENTAE STANDARD; PRT; 169 AA.
AC P08848;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division inhibitor.
GN SULA.
OS Enterobacter aerogenes (Aerobacter aerogenes).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Enterobacter.
OX NCBI_TaxID=548;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248093; PubMed=3297925;
RA Freudl R., Braun G., Honore N., Cole S.T.;
RT "Evolution of the enterobacterial sula gene: a component of the SOS
RT system encoding an inhibitor of cell division.";
RL Gene 52:31-40(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Chillingworth T., Connor P.,
RA Baker S., Basham D., Brooks R.M., Dowd L., White N., Farrar J.,
RA Cronin A., Davis N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Rellwell T., Hamlin N., Leach A., Mien T.T., Holroyd S., Parry C.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Stevens K.,
RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RA "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22591367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18";
RL J. Bacteriol. 185:2330-2337(2003).
RN [5]
RP FUNCTION: SULA IS A COMPONENT OF THE SOS SYSTEM AND AN INHIBITOR
CC OF CELL DIVISION. ACCUMULATION OF SULA CAUSES RAPID CESSATION OF
CC CELL DIVISION AND THE APPEARANCE OF LONG, NON-SEPTATE FILAMENTS.
CC THE EXPRESSION OF SULA IS REPPRESSED BY LEXA PROTEIN. FTSZ SEEMS
CC TO BE THE TARGET OF SULA.
CC -!- SUBCELLULAR LOCATION: Inner membrane.
CC -!- MISCELLANEOUS: THE N-TERMINUS OF SULA PROTEIN MIGHT BE INVOLVED IN
CC RECOGNIZING THE CELL DIVISION APPARATUS.
CC -!- SIMILARITY: THE C-TERMINUS SHOWS SIMILARITY TO THE N PROTEIN OF
CC PHAGE LAMBDA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16467; AAA24810.1; -.
DR F01667; C29016;
DR InterPro; IPR004596; Sula.
DR Pfam; PF03846; Sula; 1.
DR TIGRFAMs; TIGR00623; sula; 1.
KW Cell division; Septation; SOS response; Inner membrane.
FT DOMAIN 31 149 CONSERVED REGION.
FT DOMAIN 160 169 LON PROTEIN BINDING SITE (PROBABLE).
SQ SEQUENCE 169 AA; 19180 MW; 3099CCD8AC568A2F CRC64;

Query Match 20.9%; Score 29; DB 1; Length 169;
Best Local Similarity 35.7%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
DB 63 WQLWLTPQOKLSKE 76
```

```
RESULT 42
SULA_SALTY
ID SULA_SALTY STANDARD; PRT; 169 AA.
AC P08847;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell division inhibitor.
GN SULA OR STM1071 OR STY1092 OR T1849.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=S.typhimurium;
RX MEDLINE=87248093; PubMed=3297925;
RA Freudl R., Braun G., Honore N., Cole S.T.;
RT "Evolution of the enterobacterial sula gene: a component of the SOS
RT system encoding an inhibitor of cell division.";
RL Gene 52:31-40(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Chillingworth T., Connor P.,
RA Baker S., Basham D., Brooks R.M., Dowd L., White N., Farrar J.,
RA Cronin A., Davis N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Rellwell T., Hamlin N., Leach A., Mien T.T., Holroyd S., Parry C.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Stevens K.,
RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RA "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22591367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18";
RL J. Bacteriol. 185:2330-2337(2003).
RN [5]
RP FUNCTION: SULA IS A COMPONENT OF THE SOS SYSTEM AND AN INHIBITOR
CC OF CELL DIVISION. ACCUMULATION OF SULA CAUSES RAPID CESSATION OF
CC CELL DIVISION AND THE APPEARANCE OF LONG, NON-SEPTATE FILAMENTS.
CC THE EXPRESSION OF SULA IS REPPRESSED BY LEXA PROTEIN. FTSZ SEEMS
CC TO BE THE TARGET OF SULA.
CC -!- SUBCELLULAR LOCATION: Inner membrane.
CC -!- MISCELLANEOUS: THE N-TERMINUS OF SULA PROTEIN MIGHT BE INVOLVED IN
CC RECOGNIZING THE CELL DIVISION APPARATUS.
CC -!- SIMILARITY: THE C-TERMINUS SHOWS SIMILARITY TO THE N PROTEIN OF
CC PHAGE LAMBDA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```



```
RL J. Gen. Virol. 69:227-231 (1988).
RP SEQUENCE FROM N.A.
RA Uyeda I., Matsumura T., Sano T., Ohshima K., Shikata E.;
RT "Nucleotide sequence of rice dwarf virus genome segment 10.";
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 63:227-230 (1987).
CC -!- SIMILARITY: TO WOUND TUMOR VIRUS S10.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D00241; BAA00171.1; -
CC EMBL; M35118; AAA4732.1; -
CC InterPro; IPR008777; Phytoreo_Pns.
CC Pfam; PF05451; Phytoreo_Pns; I.
CC Nonstructural protein.
CC CONFLICT 17 17 C -> S (IN REF. 2). CRC64;
CC SEQUENCE 353 AA; 39226 MW; 8C7EC0CB559C89D0 CRC64;
CC -----
CC Query Match 20.9%; Score 29; DB 1; Length 353;
CC Best Local Similarity 23.5%; Pred. No. 37;
CC Matches 4; Conservative 2; Mismatches 11; Indels 0; Gaps 0;
CC -----
QY 39 WXXWXXXTXLNEXXL 55
DB 279 WRGWASTYAYMFNQQL 295
-----
RESULT 46
VSGP_EBOEC
ID VSGP_EBOEC STANDARD; PRT; 364 AA.
AC P87670;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Pre-small/secreted glycoprotein precursor (pre-sgp) [Contains:
DE Small/secreted glycoprotein (sgp); Delta-peptide].
GN GP.
OS Ebola virus (strain Eckron-76) (Ebo).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
OC Ebola-like viruses.
OC NCBI_TaxID=129000;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97329084; PubMed=9185597;
RA Volchkov V., Volchkova V., Eckel C., Klenk H.D., Bouloy M.,
RA Leguenno B., Feldmann H.;
RT "Emergence of subtype Zaire Ebola virus in Gabon.";
RL Virology 232:139-144 (1997).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: Pre-sgp is N-glycosylated. This precursor is processed into
CC mature sgp and delta-peptide by the host furin. Both cleavage
CC fragments contain sialic acid, but only the delta-peptide is O-
CC glycosylated (By similarity).
CC -!- RNA EDITING: Modified positions=295; Note=Partially edited. RNA
CC editing at this position consists of an insertion of one adenine
CC nucleotide (By similarity). The sequence displayed here is the
CC small secreted glycoprotein, derived from the edited RNA. The
CC unedited RNA gives rise to the full-length transmembrane
CC glycoprotein (AC P87671).
CC -!- SIMILARITY: Belongs to the filoviruses glycoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U01161; AAC57993.1; -
CC InterPro; IPR002561; Filo_glycop.
CC Pfam; PF01611; Filo_glycop; 1.
CC Glycoprotein; Signal; RNA editing.
CC SIGNAL 1 32 POTENTIAL.
CC CHAIN 33 364 PRE-SMALL/SECRETED GLYCOPROTEIN (BY
CC SIMILARITY).
CC CHAIN 33 324 SMALL/SECRETED GLYCOPROTEIN (BY
CC SIMILARITY).
CC CHAIN 325 364 DELTA PEPTIDE (BY SIMILARITY).
CC CARBOHYD 40 40 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 204 204 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 257 257 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 364 AA; 41202 MW; 5D034F8DA5EE2695 CRC64;
CC -----
CC Query Match 20.9%; Score 29; DB 1; Length 364;
CC Best Local Similarity 36.4%; Pred. No. 38;
CC Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
CC -----
QY 39 WXXWXXXTXL 49
DB 288 WAFWETKXTSL 298
-----
RESULT 47
VSGP_EBOG4
ID VSGP_EBOG4 STANDARD; PRT; 364 AA.
AC O11458;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Pre-small/secreted glycoprotein precursor (pre-sgp) [Contains:
DE Small/secreted glycoprotein (sgp); Delta-peptide].
GN GP.
OS Ebola virus (strain Gabon-94) (Ebo).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
OC Ebola-like viruses.
OC NCBI_TaxID=128947;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97329084; PubMed=9185597;
RA Volchkov V., Volchkova V., Eckel C., Klenk H.D., Bouloy M.,
RA Leguenno B., Feldmann H.;
RT "Emergence of subtype Zaire Ebola virus in Gabon.";
RL Virology 232:139-144 (1997).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: Pre-sgp is N-glycosylated. This precursor is processed into
CC mature sgp and delta-peptide by the host furin. Both cleavage
CC fragments contain sialic acid, but only the delta-peptide is O-
CC glycosylated (By similarity).
CC -!- RNA EDITING: Modified positions=295; Note=Partially edited. RNA
CC editing at this position consists of an insertion of one adenine
CC nucleotide (By similarity). The sequence displayed here is the
CC small secreted glycoprotein, derived from the edited RNA. The
CC unedited RNA gives rise to the full-length transmembrane
CC glycoprotein (AC O11457).
CC -!- SIMILARITY: Belongs to the filoviruses glycoprotein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U77384; AAC57990.1; -
CC InterPro; IPR002561; Filo_glycop.
CC -----
```



```
CC percent of the mRNA isolated from infected cells is edited. RNA
CC editing at this position consists of an insertion of one adenine
CC nucleotide. The sequence displayed here is the small secreted
CC glycoprotein, derived from the edited RNA. The unedited RNA gives
CC rise to the full-length transmembrane glycoprotein (AC Q05320).
CC
CC -!- SIMILARITY: Belongs to the filoviruses glycoprotein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U23187; AAC54886.1; -.
CC DR EMBL; U31033; AAA96745.1; -.
CC DR EMBL; AF086833; AAD14584.1; -.
CC DR InterPro; IPR002561; FiloGlycop.
CC DR Pfam; PF01611; FiloGlycop; 1.
CC KW Glycoprotein; Signal; RNA editing.
CC FT SIGNAL 1 32 POTENTIAL.
CC FT CHAIN 33 364 PRE-SMALL/SECRETED GLYCOPROTEIN.
CC FT CHAIN 33 324 SMALL/SECRETED GLYCOPROTEIN.
CC FT CHAIN 325 364 DELTA PEPTIDE.
CC FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT MUTAGEN 324 324 R->S; LOSS OF CLEAVAGE.
CC SQ SEQUENCE 364 AA; 41175 MW; 67376AA54CE5F362 CRC64;
CC
CC Query Match 20.9%; Score 29; DB 1; Length 364;
CC Best Local Similarity 36.4%; Pred. No. 38;
CC Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 39 WXXWXXXQTXL 49
CC DB 288 WAFWETKXTSL 298
CC
CC RESULT 50
CC ENO_S3YN3
CC ID ENO_S3YN3 STANDARD; PRT; 432 AA.
CC AC P77572;
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
CC glycerate hydro-lyase).
CC GN ENO OR SLR0752.
CC OS Synecocystis sp. (strain PCC 6803).
CC OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
CC OX NCBI_TaxID=1148;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=97061201; PubMed=8905231;
CC RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
CC Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
CC Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
CC Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
CC Tabata S.;
CC RT "Sequence analysis of the genome of the unicellular cyanobacterium
CC Synecocystis sp. strain PCC6803. II. Sequence determination of the
CC entire genome and assignment of potential protein-coding regions.";
CC RL DNA Res 3:109-136(1996).
CC CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC CC -!- COFACTOR: Magnesium is required for catalysis and for stabilizing
CC the dimer (By similarity).
CC CC -!- PATHWAY: Glycolysis.
```

this Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:29:11 ; Search time 39 Seconds
(without alignments)

525.864 Million cell updates/sec

Title: US-09-955-502-1

Perfect score: 139

Sequence: 1 MRRXXXCXXXXXXXXXXXXX.....QTLXNEXLXXXXXXXXXX 65

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 50 summaries

Database :

SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mmc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	33.8	94	16 Q7VK86	Q7VK86 haemophilus
2	44	31.7	90	16 Q7WH06	Q7WH06 bordetella
3	44	31.7	90	16 Q7W9Q2	Q7W9Q2 bordetella
4	44	31.7	90	16 Q7VWC4	Q7VWC4 bordetella
5	43	30.9	79	16 Q7VRC9	Q7VRC9 candidatus
6	38	27.3	634	16 Q8AB2	Q8AB2 bacteroides
7	35	25.2	1767	5 Q8W3P1	Q8W3P1 drosophila
8	34	24.5	172	16 Q831X6	Q831X6 enterococcus
9	34	24.5	221	5 Q8SS09	Q8SS09 encephalito
10	34	24.5	764	4 Q8HAP8	Q8HAP8 homo sapien
11	33	23.7	132	2 Q8EY23	Q8EY23 xanthomonas
12	33	23.7	154	16 Q8PFP4	Q8PFP4 xanthomonas
13	33	23.7	159	2 Q83RG2	Q83RG2 treponoma m
14	33	23.7	385	5 Q8VUP0	Q8VUP0 drosophila
15	33	23.7	393	16 Q8RJX2	Q8RJX2 streptomyce
16	33	23.7	393	16 Q826P2	Q826P2 streptomyce

17	33	23.7	444	5 Q9NK86	Q9NK86 drosophila
18	33	23.7	494	16 Q83KM9	Q83KM9 shigella fl
19	33	23.7	569	16 Q8FGJ7	Q8FGJ7 escherichia
20	33	23.7	1235	16 Q8Y511	Q8Y511 listeria mo
21	32	23.0	180	12 Q7TIF5	Q7TIF5 rhesus cyto
22	32	23.0	379	16 Q8DGC2	Q8DGC2 synecococc
23	32	23.0	380	12 Q9YW02	Q9YW02 melanoplus
24	32	23.0	388	11 Q8BL27	Q8BL27 mus musculu
25	32	23.0	420	16 P73615	P73615 synecocyst
26	32	23.0	518	5 Q97352	Q97352 trypanosoma
27	32	23.0	614	16 Q99F06	Q99F06 bradyrhizob
28	32	23.0	1194	4 Q9P2F6	Q9P2F6 homo sapien
29	31	22.3	232	2 Q83YT7	Q83YT7 streptococc
30	31	22.3	330	10 Q7X6M7	Q7X6M7 oryza sativ
31	31	22.3	351	10 Q8LRJ0	Q8LRJ0 oryza sativ
32	31	22.3	468	16 Q8A4I2	Q8A4I2 bacteroides
33	31	22.3	493	2 Q30495	Q30495 pseudomonas
34	31	22.3	553	16 Q7UFS4	Q7UFS4 rhodospirell
35	31	22.3	598	16 Q8XYX5	Q8XYX5 raistonia s
36	31	22.3	624	16 Q8G5P3	Q8G5P3 bifidobacte
37	31	22.3	652	2 Q93TK2	Q93TK2 streptococc
38	31	22.3	751	10 Q9AY55	Q9AY55 oryza sativ
39	31	22.3	751	10 Q7XC42	Q7XC42 oryza sativ
40	30	21.6	99	16 Q8XAH2	Q8XAH2 escherichia
41	30	21.6	99	16 Q8PDD3	Q8PDD3 escherichia
42	30	21.6	99	16 Q8XGR1	Q8XGR1 salmonella
43	30	21.6	99	16 Q83JI9	Q83JI9 shigella fl
44	30	21.6	99	16 Q7UBH5	Q7UBH5 shigella fl
45	30	21.6	124	11 Q8CEM1	Q8CEM1 mus musculu
46	30	21.6	124	13 Q818Q8	Q818Q8 gallus gall
47	30	21.6	126	16 Q8FIP4	Q8FIP4 xanthomonas
48	30	21.6	135	11 Q8BTD3	Q8BTD3 mus musculu
49	30	21.6	189	11 Q8D5B7	Q8D5B7 mus musculu
50	30	21.6	235	5 Q86AM7	Q86AM7 dictyosteli

ALIGNMENTS

RESULT 1

Q7VK86 PRELIMINARY; PRT; 94 AA.
AC Q7VK86;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
GN HD2003.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=35000HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.,;
RL "the complete genome sequence of Haemophilus ducreyi.";
DR EMBL; AE017156; AAP96716.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 94 AA; 11039 MW; B5A232C56DE93765 CRC64;

Query Match 33.8%; Score 47; DB 16; Length 94;
Best Local Similarity 23.8%; Pred. No. 0.0053;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MRRXXXCXXXXXXXXXXXXXPPXXXXXXXXXXXXXQTLXNEXL 55

Db 1 MRRVFCYLLKKEAGLDLPQYFGLGKRFNFSISKQAWAEWTKQTMLVNEKKL 55

RESULT 2


```

RA van Ham R.C.H.J., Gross R., Moya A.;
RT "The genome sequence of Blochmannia floridanus: comparative analysis
RL of reduced genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393(2003).
DR EMBL: BX248595; CAB83319.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 79 AA; 9438 MW; AE44B1E2E10FBED CRC64;

Query Match 30.9%; Score 43; DB 16; Length 79;
Best Local Similarity 30.3%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 23 PXXGXXXXXXXWXXWXXWXXQTXLXNEXXL 55
Db 23 PGKLGIVYKNISQEAANKWQNVOTILINENKL 55

RESULT 6
Q8ABB2 PRELIMINARY; PRT; 634 AA.
ID Q8ABB2;
AC 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN BT0198.
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482; ATCC 29148;
RX MEDLINE=24550858; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chang H.C., Hooper L.V., Gordon J.I.;
RA "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL: AE016926; AA075305.1; -.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0009072; F:aromatic amino acid family metabolism; IEA.
DR InterPro; IPRO01273; Aaa hydroxylase.
DR InterPro; IPRO00437; Prok lipoprotein.
DR ProDom; PD002559; Aaa hydroxylase.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 634 AA; 70192 MW; A3A4F5328ECEP469 CRC64;

Query Match 27.3%; Score 38; DB 16; Length 634;
Best Local Similarity 35.3%; Pred. No. 3.5;
Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNEXXL 55
Db 527 WNDWNMLKTLJSDASL 543

RESULT 7
Q9W3P1 PRELIMINARY; PRT; 1767 AA.
ID Q9W3P1;
AC Q9W3P1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CG15327 protein.
GN CG15327.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583; ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927;

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cui Y., Dahlke C., Davenport L.B., Davies P.,
RA dePablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hartis N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003442; AAF46280.1; -.
DR FlyBase; FBtr0029367; CG15327.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPRO01311; SBP/glu_receptor.
SQ SEQUENCE 1767 AA; 202884 MW; 243276182343EEC6 CRC64;

Query Match 25.2%; Score 35; DB 5; Length 1767;
Best Local Similarity 35.7%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNE 52
Db 1183 WSRWELSESYLRNE 1196

RESULT 8
Q831X6 PRELIMINARY; PRT; 172 AA.
ID Q831X6;
AC Q831X6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN EF2369.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583; ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927;

```

RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.P.,
RA Tetelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA Vamathevan J., Iran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA Utterback T., Racune D., Ketchum K.A., Dougherty B.A., Fraser C.M.,
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis,"
RL Science 299:2071-2074(2003).
DR EMBL: AE016954; AAC02031.1; --
DR TIGR: EF2369; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 172 AA; 20235 MW; 2F67564289B653E4 CRC64;
Query Match 24.5%; Score 34; DB 16; Length 172;
Best Local Similarity 42.9%; Pred. No. 8.2;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 39 WXXWXXXQTXLXNE 52
Db 34 WQAWKEGQFGLDNE 47
RESULT 9
Q8SS09 PRELIMINARY; PRT; 221 AA.
AC Q8SS09;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ADP ribosylation factor-like GTP binding protein.
GN EC005_0090.
OS Eucophalitozoon cuniculi.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OC NCBI_TaxID=6035;
RN [1]
RC STRAIN=GB-M1;
RP SEQUENCE FROM N.A.
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prensier G., Barbe V., Peyrecallade E., Brottier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi,"
RL Nature 414:450-453(2001).
DR EMBL: AJ590445; CAD26526.1; --
DR InterPro: IPR006689; ARF/SAR.
DR InterPro: IPR006687; SARI.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00025; arf; 1.
DR PRINTS: PR00328; SARIGTPBP.
DR SMART: SM00178; SAR; 1.
DR TIGRFAMS: TIGR00231; small GTP; 1.
SQ SEQUENCE 221 AA; 25102 MW; 6BA2DEA6B800AB98 CRC64;
Query Match 24.5%; Score 34; DB 5; Length 221;
Best Local Similarity 28.6%; Pred. No. 10;
Matches 4; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
Qy 39 WXXWXXXQTXLXNE 52
Db 160 WKSWSLQETGIENQ 173
RESULT 10
Q9HAP8 PRELIMINARY; PRT; 764 AA.
ID Q9HAP8

AC Q9HAP8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Testis development protein PRTD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Cheng L.J., Li J.M., Sha J.H.;
RT "A novel gene related to testis development (PRTD).";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF311326; AAG33852.1; --
SQ SEQUENCE 764 AA; 80380 MW; 30077783C468EE6F CRC64;
Query Match 24.5%; Score 34; DB 4; Length 764;
Best Local Similarity 35.7%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
Qy 39 WXXWXXXQTXLXNE 52
Db 17 WTGMTAQNKLFNK 30
RESULT 11
Q9EY23 PRELIMINARY; PRT; 132 AA.
ID Q9EY23;
AC Q9EY23;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative acyl carrier protein.
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=64187;
RN [1]
RP SEQUENCE FROM N.A.
RA Goel A.K., Rajagopal L., Nagesh N., Sonti R.V.;
RT "Xanthomonadin biosynthetic gene cluster from Xanthomonas oryzae pv.
RT oryzae,"
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY010120; AAG38843.1; --
DR InterPro: IPR006162; Ppantne S.
DR InterPro: IPR006163; Pp bind_
DR Pfam: PF00550; pp-binding; 1.
DR PROSITE: PS50075; ACP DOMAIN; 1.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 1.
KW Phosphopantetheine.
SQ SEQUENCE 132 AA; 14502 MW; CD7734D108E0873A CRC64;
Query Match 23.7%; Score 33; DB 2; Length 132;
Best Local Similarity 35.7%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 39 WXXWXXXQTXLXNE 52
Db 40 WKSWSLQETGIENQ 53
RESULT 12
Q8PF84 PRELIMINARY; PRT; 154 AA.
ID Q8PF84
AC Q8PF84;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Acyl carrier protein.
GN ACPC OR XAC4101.
OS Xanthomonas axonopodis (pv. citri).

OC	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC	Xanthomonadaceae; Xanthomonas.
OX	NCBI_TaxID=92829;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=306 / ATCC 13902 / XV 101;
RX	MEDLINE=22022145; PubMed=12024217;
RA	da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA	Quaggio R.B., Monteiro-Vitorlo C.B., Van Sluys M.A., Almeida N.F.,
RA	Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA	Camarotte G., Cannava F., Cardoso J., Chamargo F., Clapina L.P.,
RA	Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA	Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA	Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA	Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA	Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA	Moreira L.M., Novo M.T.M., Okura V.D., Oliveira M.C., Oliveira V.R.,
RA	Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA	Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA	Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA	Setubal J.C., Kitajima J.P.;
RT	"Comparison of the genomes of two Xanthomonas pathogens with differing
RT	host specificities."
RL	Nature 417:459-463(2002).
DR	EMBL; AE012058; AAM38936.1; -
DR	InterPro; IPR006162; Pfam; S.
DR	InterPro; IPR006163; Pfam; S.
DR	Pfam; PF00550; pp-binding; 1.
DR	PROSITE; PS00075; ACP_DOMAIN; 1.
DR	PROSITE; PS00012; PHOSPHOTRANSFERASE; 1.
KW	Complete proteins.
SQ	SEQUENCE 154 AA; 17295 MW; B30C475916875039 CRC64;
Query Match 23.7%; Score 33; DB 16; Length 154;	
Best Local Similarity 35.7%; Pred. No. 12;	
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;	
QY	39 WXXWXXQTXLXNE 52
DB	62 WKSWSQTAARE 75
RESULT 13	
ID	Q93RG2
AC	PRELIMINARY; PRT; 159 AA.
DT	01-DEC-2001 (TREMELrel. 19, Created)
DT	01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT	01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE	Flagellar protein.
GN	FLGD.
OS	Treponema medium.
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX	NCBI_TaxID=58231;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=ATCC700293;
RA	Fukunaga M.;
RT	"A phylogenetic analysis of a human oral spirochete Treponema medium
RT	by flagellar genes."
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AB046578; BAB62244.1; -
DR	InterPro; IPR005648; FLGD.
DR	Pfam; PF03963; FLGD; 1.
SQ	SEQUENCE 159 AA; 17536 MW; 8CA161F306CA8E7B CRC64;
Query Match 23.7%; Score 33; DB 2; Length 159;	
Best Local Similarity 38.5%; Pred. No. 13;	
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
QY	39 WXXWXXQTXLXN 51

DB	146 WYMSDVQTVYAN 158
RESULT 14	
Q9VJPO	PRELIMINARY; PRT; 385 AA.
ID	Q9VJPO
AC	Q9VJPO
DT	01-MAY-2000 (TREMELrel. 13, Created)
DT	01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT	01-VAR-2001 (TREMELrel. 16, Last annotation update)
DE	CG15256 protein.
GN	EG:DS04862.2 OR CG15256
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Berkley;
RX	MEDLINE=20196006; PubMed=10731132;
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Ffankoch C., Baldwin D.,
RA	Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA	Burtis K.C., Busan D.A., Butler H., Gadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Fosler C., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA	Spiel E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT	"The genome sequence of Drosophila melanogaster."
EL	Science 287:2185-2195(2000).
DR	EMBL; AB003648; ABF53466.1; -
DR	FlyBase; FBGN0028880; EG:DS04862.2.
SQ	SEQUENCE 385 AA; 45065 MW; 28933FEC7BFC6537 CRC64;
Query Match 23.7%; Score 33; DB 5; Length 385;	
Best Local Similarity 41.2%; Pred. No. 29;	
Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;	
QY	39 WXXWXXQTXLXNEXL 55
DB	199 WQHFYELQTLXNEFL 215
RESULT 15	

Q9RJX2
 ID Q9RJX2 PRELIMINARY; PRT; 393 AA.
 AC Q9RJX2;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DE Putative acyl-CoA dehydrogenase.
 GN SC01198 OR SCG11A.29C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 ON NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2); M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wierozorek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AL939108; CAB61610.1; -;
 DR GO; GO:0003995; F:acyl-CoA dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR006090; Acyl-CoA_dh_C.
 DR InterPro; IPR006091; Acyl-CoA_dh_M.
 DR Pfam; PF00441; Acyl-CoA_dh; 1_N.
 DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
 DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
 KW Complete proteome.
 SQ SEQUENCE 393 AA; 43350 MW; E3B7D6B714C6EF8F CRC64;
 Query Match 23.7%; Score 33; DB 16; Length 393;
 Best Local Similarity 54.5%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 42 WXXXQTXLXNE 52
 Db 220 WRVAQTLLNNE 230
 RESULT 16
 Q826P2
 ID Q826P2 PRELIMINARY; PRT; 393 AA.
 AC Q826P2;

DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DE Putative acyl-CoA dehydrogenase.
 GN FADEL6 OR SAV140.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 ON NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis.";
 RL Nat. Biotechnol. 21:526-531(2003).
 DR EMBL; AP005049; BAC74851.1; -;
 DR GO; GO:0003995; F:acyl-CoA dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR006090; Acyl-CoA_dh_C.
 DR InterPro; IPR006091; Acyl-CoA_dh_M.
 DR Pfam; PF00441; Acyl-CoA_dh; 1.
 DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
 KW Complete proteome.
 SQ SEQUENCE 393 AA; 43622 MW; A49E75F200B6AB4B CRC64;
 Query Match 23.7%; Score 33; DB 16; Length 393;
 Best Local Similarity 54.5%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 42 WXXXQTXLXNE 52
 Db 220 WRVAQTLLNNE 230
 RESULT 17
 Q9NK86
 ID Q9NK86 PRELIMINARY; PRT; 444 AA.
 AC Q9NK86;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DE Putative acyl-CoA dehydrogenase.
 GN BG:DS04862.2 OR CGI5256.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Sphingroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=99403001; PubMed=104711707;
 RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
 RA Celisniker S., Rubin G.M.;
 RT "An exploration of the sequence of a 2.9-Mb region of the genome of


```

RT Drosophila melanogaster: the Adh region. ";
RL [2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Celnikoff S.E., Agbayan A., Arcaina T.T., Baxter E., Blazek R.G.,
RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomocan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Shih E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE003413; AA04947.1; -.
DR FlyBase: FBgn028880; BG:DS04862.2.
KW Hypothetical protein.
SQ SEQUENCE 444 AA; 51889 MW; BAF69A2E507B92C CRC64;

Query Match 23.7%; Score 33; DB 5; Length 444;
Best Local Similarity 41.2%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLNEXXL 55
Db 258 WQHFYELQTLTYNFGAL 274

RESULT 18
Q83KW9 PRELIMINARY; PRT; 494 AA.
AC Q83KW9
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Orf, conserved hypothetical protein.
GN YEDQ OR SF2000.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL: AE015217; AAN43546.1; -.
DR InterPro: IPR00160; GGDEF; 1.
DR Pfam: PF00990; GGDEF; 1.
DR SMART: SMC0267; DUF1; 1.
DR TIGRFAMS: TIGR00254; GGDEF; 1.
DR PROSITE: PS50887; GGDEF; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 494 AA; 56208 MW; 48D653A00AFE5DF CRC64;

Query Match 23.7%; Score 33; DB 16; Length 494;
Best Local Similarity 35.3%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLNEXXL 55
Db 326 WQAWHDTLTLNFGAL 342

RESULT 19
Q8FGJ7 PRELIMINARY; PRT; 569 AA.
AC Q8FGJ7
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein yedQ.
GN YEDQ OR C2374.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Sonnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL: AE016762; AAN80833.1; -.
DR InterPro: IPR00150; GGDEF; 1.
DR Pfam: PF00990; GGDEF; 1.
DR SMART: SMC0267; DUF1; 1.
DR TIGRFAMS: TIGR00254; GGDEF; 1.
DR PROSITE: PS50887; GGDEF; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 569 AA; 64741 MW; DD1D21C8BD6E7A70 CRC64;

Query Match 23.7%; Score 33; DB 16; Length 569;
Best Local Similarity 35.3%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLNEXXL 55
Db 401 WQAWHDTLTLNFGAL 417

RESULT 20
Q8Y511 PRELIMINARY; PRT; 1235 AA.
AC Q8Y511
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein lmo2267.
GN LMO2267.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Etian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkaj G.,
RA Madueno E., Maitounan A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001)
DR EMBL: AL591982; CAD00345.1; -.
DR Pir: AC1358; AC1358.
DR Listlist; LMO02267; -.
DR GO; GO:000524; F:ATP binding; IEA.

```

DR GO: GO:0004003; F:ATP dependent DNA helicase activity; IEA.
 DR GO: GO:0003677; F:DNA binding; IEA.
 DR GO: GO:0008281; F:DNA repair; IEA.
 DR InterPro: IPR001064; Crystalilin.
 DR InterPro: IPR000212; UvrD-helicase.
 DR Pfam: PF00580; UvrD-helicase; 1.
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1235 AA; 142655 MW; E0329AD6F07E4716 CRC64;

Query Match 23.7%; Score 33; DB 16; Length 1235;
 Best Local Similarity 30.8%; Pred. No. 89;
 Matches 4; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
 Db 276 WSSWTHLXTSIEN 288

RESULT 21

Q7TF15 PRELIMINARY; PRT; 180 AA.
 AC Q7TF15
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Rhl57.
 OS Rhesus cytomegalovirus (strain 68-1) (RHCWV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=103930;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=68-1.
 RX PubMed=12767982;
 RA Hansen S.G., Strelow L.I., Franchi D.C., Anders D.G., Wong S.W.;
 RT "Complete Sequence and Genomic Analysis of Rhesus Cytomegalovirus.";
 RL J. Virol. 77:6620-6636(2003).
 DR EMBL: AY186194; AAP50679.1; -.
 SQ SEQUENCE 180 AA; 21556 MW; C9A61C4EA198DE0E CRC64;

Query Match 23.0%; Score 32; DB 12; Length 180;
 Best Local Similarity 30.8%; Pred. No. 24;
 Matches 4; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
 Db 66 WLTWHTAKTACN 78

RESULT 22

Q8DGC2 PRELIMINARY; PRT; 379 AA.
 AC Q8DGC2
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Tll2396 protein.
 GN Tll2396.
 OS Synechococcus elongatus (Thermosynechococcus elongatus).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxID=32046;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BP-1;
 RX MEDLINE=22225144; PubMed=12240834;
 RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
 RA Matarabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
 RA Shimpō S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the thermophilic cyanobacterium
 Thermosynechococcus elongatus BP-1.";
 RL DNA Res. 9:123-130(2002).

DR EMBL: AP005377; BAC09948.1; -.
 DR InterPro: IPR005338; UPF0075.
 DR Pfam: PF03702; UPF0075; 1.
 KW Complete proteome.
 SQ SEQUENCE 379 AA; 41365 MW; 639SALCADFC84975 CRC64;

Query Match 23.0%; Score 32; DB 16; Length 379;
 Best Local Similarity 30.8%; Pred. No. 49;
 Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
 Db 120 WIAWRTGNTTIAN 132

RESULT 23

Q9YW02 PRELIMINARY; PRT; 380 AA.
 AC Q9YW02
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE ORF MSV090 putative Molluscum contagiosum virus MC121L (vaccinia A16L)
 DE homolog, similar to GB:U60315.
 GN MSV090.
 OS Melanoplus sanguinipes entomopoxvirus (MsBPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
 OC Entomopoxvirus B.
 OX NCBI_TaxID=83191;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tucson;
 RX MEDLINE=99102612; PubMed=9847359;
 RA Afonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;
 RT "The genome of Melanoplus sanguinipes entomopoxvirus.";
 RL J. Virol. 73:533-552(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tucson;
 RA Afonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF063866; AAC97640.1; -.
 DR PIR: T28251; T28251.
 DR InterPro: IPR004251; DUF230.
 DR Pfam: PF03003; DUF230; 1.
 SQ SEQUENCE 380 AA; 42842 MW; FF954C5E5687C92D CRC64;

Query Match 23.0%; Score 32; DB 12; Length 380;
 Best Local Similarity 42.9%; Pred. No. 49;
 Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 42 WXXQTXLXNEXXL 55
 Db 186 WLESQTKLNNDIAL 199

RESULT 24

Q8BLZ7 PRELIMINARY; PRT; 388 AA.
 AC Q8BLZ7
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical pleckstrin homology (Fragment).
 GN A530023E23RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Aorta, and Vein;
 RX MEDLINE=22354683; PubMed=12466851;

RA The PANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK040756; BAC30694.1; -.
 DR MGI; MGI:2445175; A530023E23Rik.
 DR GO; GO:0007218; P:neuropeptide signaling pathway; IEA.
 DR InterPro; IPR001849; PH.
 DR Pfam; PF00159; PH; 1.
 DR Pfam; PF00788; RA; 1.
 DR SMART; SM00233; PH; 1.
 DR PROSITE; PS02000; RA; 1.
 KW Hypothetical protein.
 FT NON TER 388 388
 SQ SEQUENCE 388 AA; 43619 MW; 3BE39199396F511D CRC64;
 Query Match 23.0%; Score 32; DB 11; Length 388;
 Best Local Similarity 38.5%; Pred. No. 50;
 Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 39 WXXWXXXQTXLXN 51
 Db 337 WAFWGSSTHLDN 349
 RESULT 25
 ID P73615 PRELIMINARY; PRT; 420 AA.
 AC P73615;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein slr1865.
 GN SLR1865.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions."
 RL DNA Res. 3:109-136 (1996).
 DR EMBL; D90908; BAAL7660.1; -.
 DR PIR; S77102; S77102.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR006059; SBP_bac_1.
 DR Pfam; PF01547; SBP_bac_1_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 420 AA; 47851 MW; D617A38D98E8E958C CRC64;
 Query Match 23.0%; Score 32; DB 16; Length 420;
 Best Local Similarity 35.7%; Pred. No. 54;
 Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
 QY 39 WXXWXXXQTXLXNE 52
 Db 150 WRFWEVAQTNIQK 163
 RESULT 26
 ID O97352 PRELIMINARY; PRT; 518 AA.
 AC O97352;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ILTat 1.61 metacyclic VSG protein.
 OS Trypanosoma brucei.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20017521; PubMed=10551359;
 RA Graham S.V., Terry S., Barry J.D.;
 RT "A structural and transcription pattern for variant surface
 RT glycoprotein gene expression sites used in metacyclic stage
 RT Trypanosoma brucei."
 RL Mol. Biochem. Parasitol. 103:141-154 (1999).
 RN 2;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91081307; PubMed=2175429;
 RA Matthews K.R., Shields P.G., Graham S.V., Cowan C., Barry J.D.;
 RT "Duplicative activation mechanisms of two trypanosome telomeric VSG
 RT genes with structurally simple 5' flanks."
 RL Nucleic Acids Res. 18:7219-7227 (1990).
 DR EMBL; AJ012199; CAA0956.1; -.
 SQ SEQUENCE 518 AA; 55736 MW; 800D002074229468 CRC64;
 Query Match 23.0%; Score 32; DB 5; Length 518;
 Best Local Similarity 35.3%; Pred. No. 66;
 Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 39 WXXWXXXQTXLXNXXL 55
 Db 108 WEEKNKNSATRLKEAVL 124
 RESULT 27
 ID Q89FU6 PRELIMINARY; PRT; 614 AA.
 AC Q89FU6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE BL16704 protein.
 GN BL16704
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX STRAIN=USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
 RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 RT Bradyrhizobium japonicum USDA110."
 RL DNA Res. 9:189-197 (2002).
 DR EMBL; AP005959; BAC51969.1; -.
 DR GO; GO:000155; P:two-component sensor molecule activity; IEA.
 DR GO; GO:000160; P:two-component signal transduction system (p. .); IEA.
 DR InterPro; IPR000160; GGDEF.
 DR InterPro; IPR001610; PAC.
 DR InterPro; IPR00700; PAS-assoc C.
 DR InterPro; IPR000014; PAS_domain.
 DR Pfam; PF00990; GGDEF; 1.
 DR Pfam; PF00785; PAC; 1.
 DR Pfam; PF00989; PAS; 1.
 DR PROSITE; PS0887; GGDEF; 1.
 DR PROSITE; PS0113; PAC; 1.
 DR PROSITE; PS0112; PAS; 1.
 KW Complete proteome.
 SQ SEQUENCE 614 AA; 67212 MW; 1CBE3A89D5B98853 CRC64;

Query Match 23.0%; Score 32; DB 16; Length 614;
 Best Local Similarity 45.5%; Pred. No. 77;
 Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 49
 DB 14 WRENSRQTL 24

RESULT 28

Q92P26 PRELIMINARY; PRT; 1194 AA.
 AC Q92P26;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE Hypothetical protein KIAA1391 (Fragment).
 GN KIAA1391.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Brain;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes.XVI.
 RT The complete sequences of 150 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 7:65-73(2000).
 DR EMBL; AB037812; BAA92629.1; -.
 DR PIR; C59436; C59436.
 DR HSSP; Q07960; IRGP.
 DR GO; GO:0007218; P:neuropeptide signaling pathway; IEA.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000159; RA domain.
 DR InterPro; IPR000198; RhoGAP.
 DR InterPro; IPR008936; RhoGAP.
 DR Pfam; PF00788; RA; 1.
 DR Pfam; PF00620; RhoGAP; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00324; RhoGAP; 1.
 DR PROSITE; PS0200; RA; 1.
 DR PROSITE; PS0238; RHO GAP; 1.
 KW Hypothetical protein.
 FT NON-TER
 SQ SEQUENCE 1194 AA; 133015 MW; C4BB7B7F4CC9F59 CRC64;

Query Match 23.0%; Score 32; DB 4; Length 1194;
 Best Local Similarity 38.5%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
 DB 340 WAFWRGSGSTHLDN 352

RESULT 29

Q83Y77 PRELIMINARY; PRT; 232 AA.
 AC Q83Y77;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DE Amylase-binding protein B (Fragment).
 GN ABPB.
 OS Streptococcus gordonii.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1302;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=38;
 RA Xu D.-Q., Thompson J., Cisar J.O.;
 RT "Genetic Loci for Coaggregation Receptor Polysaccharide Biosynthesis
 RT in Streptococcus gordonii 38."
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY147912; AAN52126.1; -.
 FT NON-TER
 SQ SEQUENCE 232 AA; 26080 MW; 52FFA0612E7DA1FC CRC64;

Query Match 22.3%; Score 31; DB 2; Length 232;
 Best Local Similarity 28.6%; Pred. No. 52;

Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 52
 DB 2 WTVWHIDQMAIKNQ 15

RESULT 30

Q7X6M7 PRELIMINARY; PRT; 330 AA.
 AC Q7X6M7;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DE OSJNB0043L24.5 protein (OSJNB0002J11.13 protein).
 GN OSJNB0043L24.5 OR OSJNB0002J11.13.
 OS Oryza sativa (rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
 RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
 RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
 RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
 RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.

RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
 RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
 RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL662969; CAE04717.1; -.
 DR EMBL; AL731626; CAE05686.1; -.
 SQ SEQUENCE 330 AA; 35314 MW; C564E5CD33008AFF CRC64;

Query Match 22.3%; Score 31; DB 10; Length 330;
 Best Local Similarity 38.5%; Pred. No. 72;

Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
 DB 265 WAKWLDQKLAN 277

RESULT 31

Q8LRJ0 PRELIMINARY; PRT; 351 AA.
 ID Q8LRJ0

```

AC Q8LRJ0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE OSJNBa0016109.11 protein.
GN OSJNBa0016109.11
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
  clone:OSJNBa0016109."
RL Submitted (DGC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003052; BAB92159.1; -.
DR Gramene; Q8LRJ0; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 2.
SQ SEQUENCE 351 AA; 7779AD2D12386FD CRC64;

Query Match 22.3%; Score 31; DB 10; Length 351;
Best Local Similarity 35.7%; Pred. No. 76;
Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXQTXLXN 52
DB 173 WSLWVLMSPLLNE 186

RESULT 32
Q8A412 PRELIMINARY; PRT; 468 AA.
AC Q8A412;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Reverse transcriptase.
GN BT2617.
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
  Chang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."
RL Science 299:2074-2076(2003).
DR EMBL; AE016936; AA077724.1; -.
DR GO; GO:0005945; C:6-phosphofructokinase complex; IEA.
DR GO; GO:0003872; F:6-phosphofructokinase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003944; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006036; F:glycolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR000023; Pfprckinase.
DR InterPro; IPR00123; RNA DNAPolys.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00078; rvt; 1.
DR PRINTS; PR00866; RNADNAPOLMS.
DR ProDom; PD000707; Pfprckinase; 1.
KW Complete proteome.
SQ SEQUENCE 468 AA; 54642 MW; 097FCD5D4223C618 CRC64;

Query Match 22.3%; Score 31; DB 16; Length 468;
Best Local Similarity 30.8%; Pred. No. 1e-02;

```

```

Matches 4; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXQTXLXN 51
DB 399 WKSWKNAKTRVAN 411

RESULT 33
Q30495 PRELIMINARY; PRT; 493 AA.
AC Q30495;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Xylulose kinase (EC 2.7.1.17) (Xylulokinase).
GN MTLY OR XYLX.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN=DSM 50106;
RX MEDLINE=98121321; PubMed=9461423;
RA Buerker P., Altenbuchner J., Mates R.;
RT "Structure and function of the genes involved in mannitol, arabinol
  and glucitol utilization from Pseudomonas fluorescens DSM50106."
RL Gene 206:117-126(1998).
CC -1- CATALYTIC ACTIVITY: ATP + D-XYLULOSE = ADP + D-XYLULOSE 5-
  PHOSPHATE.
CC -1- INDUCTION: BY MANNITOL, ARABITOL AND SORBITOL. NOT INDUCED BY
  FRUCTOSE OR XYLULOSE.
CC -1- MISCELLANEOUS: THIS ENZYME IS SPECIFIC FOR XYLULOSE.
CC -1- MISCELLANEOUS: THE MTLY PROTEIN IS ENCODED BY THE MTL EFGKDYZ
  OPERON. THIS OPERON ENCODES PROTEINS FOR THE UPTAKE AND
  UTILIZATION OF MANNITOL, ARABITOL AND SORBITOL.
CC -1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCONOKINASE /
  GLYCEROKINASE / XYLULOKINASE FAMILY.
CC EMBL; AF007800; AAC04473.1; -.
DR GO; GO:0016301; P:kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0004856; P:xylulokinase activity; IEA.
DR GO; GO:0005996; P:monosaccharide metabolism; IEA.
DR GO; GO:0005997; P:xylulose metabolism; IEA.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR001680; WD40.
DR InterPro; IPR006000; Xylulokinase.
DR Pfam; PF0370; FGGY; 1.
DR Pfam; PF02782; FGGY_C; 1.
DR TIGRfam; TIGR01312; xylB; 1.
DR PROSITE; PS00933; FGGY_KINASES_1; FALSE_NEG.
DR PROSITE; PS00445; FGGY_KINASES_2; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR TRANSFERASE; Kinase; Xylose metabolism.
SQ SEQUENCE 493 AA; 52057 MW; 6566792F8E5B5C65 CRC64;

Query Match 22.3%; Score 31; DB 2; Length 493;
Best Local Similarity 42.9%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 42 WXXWXXQTXLXNEXXL 55
DB 104 WCDTETALENERLL 117

RESULT 34
Q7UF54 PRELIMINARY; PRT; 553 AA.
AC Q7UF54;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Protein kinase homolog (EC 2.7.1.-).

```



```

RESULT 38
Q9AY55 PRELIMINARY; PRT; 751 AA.
ID Q9AY55
AC Q9AY55;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative arm repeat protein.
GN OSJNBA0027P10.11.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Hsiao J.,
RA Zismann V., Pal G., Bowman C.L., Fujii C.F., VanAken S.E., Craven B.,
RA Khalak H., Feldblyum T.V., Quackenbush J., White O., Salzberg S.L.,
RA Fraser C.M.;
RA "Oryza sativa chromosome 10 BAC OSJNBA0027P10 genomic sequence.";
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AC084763; AAG60190.1; -.
DR HSP: P35222; IG3J.
DR Gramene; Q9AY55.
DR InterPro; IPR008936; ARM.
DR InterPro; IPR000225; Armadillo.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR007089; LRR_cys.
DR InterPro; IPR008945; Skp1_Skp2.
DR Pfam; PF00514; Armadillo_seg; 6.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00185; ARM; 6.
DR SMART; SM00256; FBOX; 1.
DR PROSITE; PSS0181; FBOX; 1.
SQ SEQUENCE 751 AA; 79515 MW; 2E33B5FEFAA26685 CRC64;

Query Match 22.3%; Score 31; DB 10; Length 751;
Best Local Similarity 45.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXL 49
Db 375 WLEWILSQTL 385

RESULT 39
Q9AY55 PRELIMINARY; PRT; 751 AA.
ID Q9AY55
AC Q9AY55;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative arm repeat protein.
GN OSJNBA0027P10.11.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA The Rice Chromosome 10 Sequencing Consortium;
RA "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";
RL Science 300:1566-1569(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;

Query Match 21.6%; Score 30; DB 16; Length 99;
Best Local Similarity 28.6%; Pred. No. 39;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXL 52
Db 82 WSAWRLVKTLKQ 95

RESULT 41
Q8FDD3 PRELIMINARY; PRT; 99 AA.
ID Q8FDD3
AC Q8FDD3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein yqjk.

```

```

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE017120; AAP55033.1; -. 2B33B5FEFAA26685 CRC64;
SQ SEQUENCE 751 AA; 79514 MW; 2E33B5FEFAA26685 CRC64;

Query Match 22.3%; Score 31; DB 10; Length 751;
Best Local Similarity 45.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXL 49
Db 375 WLEWILSQTL 385

RESULT 40
Q8XAH2 PRELIMINARY; PRT; 99 AA.
ID Q8XAH2
AC Q8XAH2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Orf, hypothetical protein.
GN Z4454 OR ECS3982.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Gotsfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Gotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RT Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:111-22(2001).
DR EMBL: AE005539; AAG58233.1; -.
DR EMBL: AP002564; BAB37405.1; -.
DR PIR; E85971; E85971.
DR PIR; F91126; F91126.
KW Complete proteome.
SQ SEQUENCE 99 AA; 11781 MW; C6B9170493DF4086 CRC64;

Query Match 21.6%; Score 30; DB 16; Length 99;
Best Local Similarity 28.6%; Pred. No. 39;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXL 52
Db 82 WSAWRLVKTLKQ 95

RESULT 41
Q8FDD3 PRELIMINARY; PRT; 99 AA.
ID Q8FDD3
AC Q8FDD3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein yqjk.

```

GN YQJK OR C3858.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=06:HI / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
 DR EMBL; AB016767; AAN82303.1; -;
 DR FIR; F91126; F85971.
 DR PIR; F91126; F85971.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 99 AA; 11781 MW; C6B9170493DF4086 CRC64;
 Query Match 21.6%; Score 30; DB 16; Length 99;
 Best Local Similarity 28.6%; Pred. No. 39;
 Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 39 WXXWXXXTLXNE 52
 Db 82 WSAWRLVKTTLKQQ 95
 RESULT 42
 Q8XGR1 PRELIMINARY; PRT; 99 AA.
 ID Q8XGR1
 AC Q8XGR1
 DT 01-MAR-2002 (TREMELrel. 20, Created)
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Hypothetical protein (Putative inner membrane protein).
 GN T3151 OR YQJK OR STG231 OR STY3411.
 OS Salmonella typhimurium.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OC NCBI_TaxID=601, 602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 and CT18.";
 RL J. Bacteriol. 185:2330-2337 (2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856 (2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Cornerton P.,

RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi CT18.";
 RL Nature 413:848-852 (2001).
 DR EMBL; AE016844; AA070692.1; -;
 DR EMBL; AE008848; AAL22104.1; -;
 DR EMBL; AL627278; CAD07755.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 99 AA; 11548 MW; 28F4194AE256CE35 CRC64;
 Query Match 21.6%; Score 30; DB 16; Length 99;
 Best Local Similarity 28.6%; Pred. No. 39;
 Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 39 WXXWXXXTLXNE 52
 Db 82 WSAWRLVKTTLKQQ 95
 RESULT 43
 Q83J19 PRELIMINARY; PRT; 99 AA.
 ID Q83J19
 AC Q83J19
 DT 01-JUN-2003 (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Orf, conserved hypothetical protein.
 GN SF3143.
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OC NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hsu Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441 (2002).
 DR EMBL; AE015325; AAN44614.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 99 AA; 11781 MW; 33B9020486DF4137 CRC64;
 Query Match 21.6%; Score 30; DB 16; Length 99;
 Best Local Similarity 28.8%; Pred. No. 39;
 Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 39 WXXWXXXTLXNE 52
 Db 82 WSAWRLVKTTLKQQ 95
 RESULT 44
 Q7UBH5 PRELIMINARY; PRT; 99 AA.
 ID Q7UBH5
 AC Q7UBH5
 DT 01-OCT-2003 (TREMELrel. 25, Created)
 DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN S351.
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.


```

OX NCBI_TaxID=623;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE016989; AAP18428.1; -.
KW Hypothetical protein.
SQ SEQUENCE 99 AA; 11781 MW; 33B90651D7DB1466 CRC64;

Query Match 21.6%; Score 30; DB 16; Length 99;
Best Local Similarity 28.6%; Pred. No. 39;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNE 52
| | | | |
Db 82 WSAWLVKTKLKQ 95
| | | | |

RESULT 45
Q8CEM1 ID Q8CEM1 PRELIMINARY; PRT; 124 AA.
AC Q8CEM1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
GN 332140IG04RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
EL Nature 420:563-573(2002).
DR EMBL; AK019001; BAC25574.1; -.
DR MGD; MGI:1914665; 332140IG04RIK.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 124 AA; 14796 MW; 4C123AEFD198E87B CRC64;

Query Match 21.6%; Score 30; DB 11; Length 124;
Best Local Similarity 36.4%; Pred. No. 48;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXL 49
| | | | |
Db 26 WNAWTALETYL 36
| | | | |

RESULT 46
Q918Q8 ID Q918Q8 PRELIMINARY; PRT; 124 AA.
AC Q918Q8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Bone morphogenetic protein 7 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;

```

```

OC NCBI_TaxID=9031;
OX [1];
RN SEQUENCE FROM N.A.
RX MEDLINE=20307867; PubMed=10848589;
RA Monroe D.G., Jin D.F., Sanders M.M.;
RT "Estrogen Opposes the Apoptotic Effects of Bone Morphogenetic Protein
RT 7 on Tissue Remodeling.";
RL Mol. Cell. Biol. 20:4626-4634(2000).
DR EMBL; AF223970; AAF34758.1; -.
DR InterPro; IPR001111; TGFb.N.
DR Pfam; PF00688; TGFb_propeptide; 1.
FT NON TER 124
SQ SEQUENCE 124 AA; 13794 MW; F70104D9196A4248 CRC64;

Query Match 21.6%; Score 30; DB 13; Length 124;
Best Local Similarity 35.7%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXNE 52
| | | | |
Db 23 WLRWLADFTLDNE 36
| | | | |

RESULT 47
Q8PIP4 ID Q8PIP4 PRELIMINARY; PRT; 126 AA.
AC Q8PIP4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein XAC2851.
GN XAC2851.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE011926; AAM37696.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 126 AA; 13919 MW; E655BB90248FDB5C CRC64;

Query Match 21.6%; Score 30; DB 16; Length 126;
Best Local Similarity 38.5%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXLXN 51
| | | | |
Db 108 WLLWQQPQTRVMN 120
| | | | |

RESULT 48

```

Q8BTD3
ID Q8BTD3 PRELIMINARY; PRT; 135 AA.
AC Q8BTD3
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
GN 3321401G04RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002)."
RL Nature 420:563-573 (2002).
DR EMBL; AK004097; BAC25066.1; --
DR MGD; MGI:1914665; 3321401G04RIK.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 135 AA; 15976 MW; CA485C37EDD3300F CRC64;
Query Match 21.6%; Score 30; DB 11; Length 135;
Best Local Similarity 36.4%; Pred. No. 52;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 39 WXXWXXQXPL 49
DB 37 WNAWALETYL 47
RESULT 49
Q9DSB7
ID Q9DSB7 PRELIMINARY; PRT; 189 AA.
AC Q9DSB7
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE 4930469G21RIK protein.
GN 4930469G21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690 (2001).

DR EMBL; AK015528; BAB29884.1; --
DR MGD; MGI:1922183; 4930469G21RIK.
SQ SEQUENCE 189 AA; 21841 MW; C445DF77CDEE361B CRC64;
Query Match 21.6%; Score 30; DB 11; Length 189;
Best Local Similarity 29.4%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;
QY 39 WXXWXXQTXLXNEXL 55
DB 99 WKKWTRHOKLKQOASL 115
RESULT 50
Q86AM7
ID Q86AM7 PRELIMINARY; PRT; 235 AA.
AC Q86AM7
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to Dictyostelium discoideum (Slime mold). prenyl cysteine
DE carboxyl methyltransferase.
DE Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.,
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum."
RL Nature 418:79-85 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116984; AA051329.1; --
DR GO; GO:0008168; F:methyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR007269; ICMT.
DR Pfam; PF04140; ICMT; 1.
KW Methyltransferase; Transferase.
SQ SEQUENCE 235 AA; 27177 MW; CCDCAA6C8AF85014 CRC64;
Query Match 21.6%; Score 30; DB 5; Length 235;
Best Local Similarity 26.6%; Pred. No. 88;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
QY 39 WXXWXXQTXLXNE 52
DB 195 WASWSFFSQRIENE 208
Search completed: September 30, 2004, 10:32:34
Job time : 43 secs